Johann Friedrich Herbart (1776–1841) is known today mainly as a founding figure of modern psychology and educational theory. But these were only parts of a much grander philosophical project, and it was as a philosopher of the first rank that his contemporaries saw him. Even in his own day, Herbart's direct influence on academic philosophy was limited, but this had as much to do with shifting disciplinary borders as with his polemics against the German Idealists. In psychology and pedagogy, however, his influence was greater and longer lasting. While no one took over his philosophy or psychology (and especially the impenetrable mathematics) as a whole, certain aspects of his thought proved immensely fruitful. Indeed, without Herbart, the landscape of modern psychology and philosophy would be unrecognizable.

Though his mathematization of the mind proved a dead end, it encouraged early experimentalists like Fechner to apply mathematics to the psyche. His interpretation of space had an important influence on Helmholtz's empirical theory of vision, and on Natorp's "logicization" of Kantian spatio-temporal intuition. Herbart's arguments against psychological faculties were taken over by Wundt, and his notions of the *limen* and the "subconscious" were fundamental to Freud's psychoanalytic theory. In philosophy, Herbart anticipated key ideas of Phenomenology, such as Brentano's conception of mental phenomena, and his ideas of representational force, the *limen*, and reproduction all find positive endorsement and development in Husserl's Phenomenology. Finally, recent scholarship in both psychology and philosophy is rediscovering Herbart. For example, Boudewijnse et al. argue not only for the contemporary relevance of Herbart's psycho-mechanics, but also that it is supported by experiment. And Beiser's important contribution to the early history of neo-Kantianism reveals Herbart as a radical and important interpreter of Kant.

1. Biographical sketch

Herbart was born on May 4, 1776, in Oldenburg.[1] He displayed precocious philosophical talent, writing on the philosophy of Wolff and Kant in his early teens. His early education emphasized music, which would be reflected in his later writings.[2] His first teacher was his mother, who home-schooled him for many years, and later followed him to Jena, where he matriculated in 1794. There he met Fichte, becoming a member of his inner circle. Already by 1794, however, Herbart took a critical stance towards the *Wissenschaftslehre*.[3] Still, it was Fichte who set the general problem of Herbart's life work: the ego and self-consciousness (cf. James 1890, Vol. I: 353–4; Beiser 2014: 101).

While working as a private tutor to a Swiss family from 1796 to 1800, Herbart met Pestalozzi, the great Swiss educator and theorist of pedagogy. Pestalozzi's critique of traditional instructional means and ends, and his emphasis on developing the pupil's autonomy, would permanently stamp Herbart's own pedagogic thought (Beiser 2014: 101). During this period, Herbart struggled to find his way out of Fichte's shadow and discover his own system (Beiser 2014: 102–3). After personal trials involving his mother, Herbart moved to Göttingen, where he passed his doctoral and habilitation exams, lecturing there from 1802–1804. In his inaugural lecture on Plato's theory of forms Herbart first mentions his "method of relations", the logical basis for his metaphysical thought (see Section 2, below). While at Göttingen, he published the major works, *Allgemeine Pädagogik* (*General Pedagogy*, 1806a), *Hauptpuncte der Metaphysik* (*Basic Points of Metaphysics*, 1806b); and *Allgemeine practische Philosophie*, 1808). When Göttingen came under French occupation in 1806, Herbart's situation there grew increasingly difficult.

He therefore gladly accepted an offer from Königsberg to become the second successor to Kant's chair in 1809 (Beiser 2014: 130–1). From his outpost in East Prussia, he watched the triumph of German Idealism and Romanticism with despair (Beiser 2014: 93; see the preface to his [1822; SW V: 94]). His 1814 essay, with the acid title, "Concerning My Quarrel with the Philosophy *du Jour* [*Ueber Meinen Streit mit der Modephilosophie dieser Zeit*]", attacks the philosophy of Schelling and its reliance on intellectual intuition (SW III: 317–52; cf. Beiser 2014: 132–4). It was during his Königsberg years that Herbart developed his psychology in works like "Psychological Remarks on the Doctrine of Tones" (1811), the *Lehrbuch zur Psychologie* (*Teaching Text for Psychology*; 1816), and his monumental *Psychologie als Wissenschaft* (*Psychology as Science*; two volumes, 1824, 1825). In 1833, Herbart returned to Göttingen, where he resided until his death on August 14, 1841.

2. Logic and metaphysics

Aside from Herbart's textbook presentation of traditional logic (e.g., Section II of LEP), his main logical innovation is his so-called method of relations, devised as an organon for correct metaphysical

thinking (SW V: 201, f.). The metaphysical problems it aims to address include: the thing as (one) substance with (many) characteristics; causality; matter; and the ego. He considers these problematic because they give rise to an empirically given, yet self-contradictory concept (Weiss 1928: 37). Due to space constraints, we will here only briefly touch on the problems of thing and ego, the former for its role in the development of apperception, the latter because it is grounds his psychology. Both his idea of metaphysics and the method of relations will become clearer as we see them used to explain concrete psychological phenomena in Section 3.

The problems of both thing and ego may be stated thus: there is a conflict between an alleged unity, on the one hand, with an empirically manifest plurality, on the other. We consider a thing to be one substance, yet possessing a number of conflicting qualities; and we regard the self as a single entity conscious of many thoughts. But *is* the thing the one, or its many qualities? *Is* the self or psyche the one, or its thoughts? Herbart maintains that this dilemma arises due to a missing link, one that can only be supplied by "speculation". Speculative use of the method of relations would have us consider the conflicting elements not as distinct entities in genuine opposition, but rather as members of a unifying relation. For example, the "thing" turns out to be the *nexus* or system of (its) predicates, and nothing else beyond them (cf. Beiser 2014: 118). Similarly, Herbart regards the psyche not as a substance, but simply as the condition of mental change itself.

He follows Kant in distinguishing between conscious phenomena and unknowable noumena that must be conceived as underwriting the coherence that *experience* finds in the fluctuating phenomenal manifold. The noumena are "speculative", by which Herbart does not mean objects of intellectual intuition, but rather the transcendental conditions of the synthetic unity of objects of experience. Yet where Kant held noumena to be unknowable, Herbart takes the further step of ascribing to every existent "an act of self-preservation", a notion as opaque as it is important to his psychology (SW II: 195). The idea seems to be that every being must have an inherent force holding it together, or there would be no reason to suppose its continued existence: "the idea of a force to maintain one's identity" just is "self-preservation" (Beiser 2014: 120).

In short, the method of relations is a way of doing metaphysics with strong Platonic and Kantian overtones: the fluctuating and conflicting plurality of sensible experience is "related" back to intelligible, "speculative" unities that allow us to organize the plurality in experience. Thus, the method does more than merely analyze contradiction; it also resolves them by supplying the intelligible links. We will take these points up in more detail as they arise in Herbart's psychology.

3. Psychology

3.1 Introduction

The main motive of speculative thinking, then, is the resolution of apparent contradictions continuously arising in experience. In particular, such contradictions arise when essentially opposed concepts are thought together as unities. This problem of the unity of opposites also arises in the case of the psyche, which is at once a unity and multiplicity: for it is one consciousness of innumerable representations (SW V: 307). Hence it is psychology's task to explain how the ego can arise as one out of a multiplicity of conjoined or conflicting inner states.

3.1.1 Critique of faculty psychology

The metaphysical assumption of the soul's simplicity immediately puts Herbart at odds with the traditional "psychology of faculties", according to which mental phenomena and conscious processes are derived from "special[ized] powers of the psyche" (Weiss 1928: 69; cf. esp. PsW: 11, ff.; Beiser 2014: 136, ff.).[4] He attacks this view, held prominently by Kant, with a battery of arguments. First, he claims that these so-called faculties are really nothing more than class concepts (PsW: 16). While we may tend to ascribe individual instances of recall to a faculty called "memory", this term is just a catchall for the instances. Rather than a "power", it is an abstraction, and hence no explanation of the phenomena from which it has been abstracted (PsW: 25, ff.; cf. LPs [1850]: 8–9). Second, the theory fails to account for the

> causal relation among the psychic faculties, through which, acting in concert, they interact and mutually prompt or induce or force each other to act. (PsW: 28 [=SW V: 199])

Third, psychic powers seem engaged in a war of all against all, forcing us to presume an irremediably fractured self (PsW: 28; 29; cf. 131). Herbart's fourth and gravest complaint is that the theory of faculties ignores the particular, and neglects exact description, the *sine qua non* of empirical science (PsW: 27).

3.1.2 Psychology as a future science

He therefore proposes to construct a new system,[5] analogous to physics (*Naturforschung*: PsW: 10), to avoid the pitfalls and disprove the theory of faculties (LPs [1850]: 5).[6] The new psychology must, like the natural sciences, assume that all phenomena are governed without exception by a set of laws, which scientists seek out through the establishment of the facts; prudent conclusions; tested and corrected hypotheses (LPs [1850]: 9); and, wherever possible, through the measurement of

magnitudes and mathematical calculation (PsW: 10). Similarly, Herbart argues, our mental representations are subject to a system of laws that let us discern the regularities of their interactions (cf. SW V: 195).

3.1.3 Critique of introspection

To discover the laws, we must first identify and exactly describe individual mental representations without recourse to contrived faculties, for representations are the sole elements of our actual mental life (PsW: 16). He identifies three possible ways of identifying these principles of consciousness:

1. self-observation (either intentional or inadvertent) (PsW: 18, 19);
2. interpretation of the products of my own activity;
3. testimony and observation of others (PsW: 10; cf. Section I of the Introduction [§§1–6]).

Each of these paths has its drawbacks. (1) Intentional or self-conscious introspection interferes in the natural operation of psychic processes (LPs [1850]: 8). Even if introspection happens incidentally, Herbart points out that the observer always already introduces some prior knowledge of himself that colors the resulting observation (SW V: 192; cf. [1840]: 36). (2) Interpretation of the products of my own activity obviously cannot serve to reveal the principles, since the products are *ex hypothesi* already removed from those principles. Rather, the products of conscious activity are the very phenomena of which our science must seek the causes (PsW: 20–21; SW V: 194). Finally, (3) the interpretation of others is inadequate because in addition to its inherent unreliability, even the most faithful testimony again depends on the *informant's* introspection. Moreover, the psychologist's interpretation of that testimony will require comparison with his own knowledge of his inner conditions, so that the drawbacks of introspection make themselves felt from two sides, not just one (SW V: 198).[7]

3.1.4 Metaphysics and the method of relations

The aforementioned ways of capturing the prime facts of psychology are all defective; no true science can content itself with such inadequate material. So Herbart turns to his method of relations. Only insofar as a particular representation may be picked out of the "whole mass of inner perceptions" (PsW: 33), in order to serve as a known relatum, standing in relation to some other, as yet unknown relatum—i.e., its condition of possibility—only to that extent does it function as a principle of psychology (PsW: 15).[8] Psychology, if it is to be scientific, must supplement and *complete* the empirical material (as such but a wild flux) by way of a principle of determinate stability: psychology aims to "prove—by means of what perception cannot attain—the nexus of what has let itself be perceived" (PsW: 32; cf. PsW: 131; Weiss 1928: 72–3). But thinking the empirically given with the aid of something hidden is only possible through metaphysics, which explains *all* being (*Seiendes*) (Hartenstein 1850: viii–ix; cf. esp. PsW: 32–3). Only in this way can empirical psychic facts be brought under a universal set of laws and overcome their isolation from physical phenomena (Weiss 1928: 73). Universal metaphysical principles, valid across the psychophysical schism, provide the framework within which the psychic facts can be "completed" and become proper objects of scientific understanding.

3.1.5 Calculation

The paradigm of physics suggests that phenomenal laws can be gained through experiment, instrumentation (measurement), and calculation. But of these three paths only calculation is available to psychology.[9] Calculation alone can reveal the basic laws and concepts of the psyche; psychic phenomena are then to be traced back or "reduced" to these rules; and in this way an "interconnected order" of the psychic manifold may be constructed, analogous to the one physics provides to external, sensible phenomena. As he puts it in "On Analogies, with Respect to the Foundation of Psychology",

> mathematics [is] not applied to concepts that come from naked, immediate experience in the manner of empiricism, but from experience *as processed through metaphysics*, and which [concepts, then,] returns, aided by mathematics, to experience. (Herbart 1840b: 185)

Thus, the traditional highest "generic concepts" of representation, feeling, and desire that were formulated via an unscientific and chaotic process of abstraction, must be replaced by whatever rules calculation may reveal (LPs [1850]: 8).

To sum up Herbart's approach: all mental life manifests itself in constant temporal flux and continuous mutation, as "a manifold of variegated determinations in one [*in Einem*]", and ultimately, as self-consciousness (LPs [1850]: 11–12; cf. PsW: 30). We must therefore begin with general metaphysical or "speculative" postulates to compensate for the obstacles to pure observation of psychic life, and so postulate the simplicity of the soul (cf. Beiser 2014: 138). Beyond this, however, psychology must avail itself of mathematics, in that

> representations must be viewed as forces, whose
> effective power depends on their strength [or intensity],
> their oppositions, and their connections, all of which
> differ in degree. (LPs [1850]: 12)

We now turn to the heart of Herbart's psychology, namely just this "force-interpretation" of mental representations, and the mathematical manipulation it makes possible.

3.2 Rules of Representations

It is Herbart's radical contention that consciousness is nothing more than the continuous stream of representations. [10] These representations, he says cryptically, are the "self-preservations of the soul [*Selbsterhaltungen der Seele*]" (PsW: 171). That is, as we saw, because the soul (the terms, "psyche" and "soul", will be used interchangeably in this article) is postulated to be simple and "originally an utter *tabula rasa*, without any life or thought" (SW VI: 120), the evident multiplicity of conscious representations generates a paradox (PsW: 171). Since the soul is by nature simple, the source of any psychic differentiation must be sought outside of it, viz., among the "objects" that affect it through the senses (PsW: 132). Thus, on the one hand, the world impinges on and perturbs the soul's natural placidity (cf., e.g., the *Māṇḍūkya* Upaniṣad; PsW: 132–3; but see Felsch 1904: 196–7); on the other hand, the soul resists these disturbances (i.e., the representations themselves) by constantly seeking to reestablish a harmonious unity or equilibrium among them and thus "preserve", as far as possible, its own simplicity.

Psychology's task, then, is to explain how the various representations mutually resolve each other to attain maximal unity (Weiss 1928: 73). They do this in two ways. First, wherever there are several representations of the same kind, they "merge" into a single representational activity ("*... Ein intensives Thun*", PsW: 171; Herbart 1811; Boudewijnse, et al. 1999: 170). On the other hand, if representations are of different or opposed kinds, they "inhibit" each other "quantitatively" (PsW: 129–30). Simply put, two representations present in consciousness may either resemble each other or not. If the former, then they will merge into one, stronger representation, and unity will be attained in this way. If they differ, then they will struggle against each other, each trying to negate the other as far as possible. In the latter case, the representing activity remains constant, its essential simplicity expressing itself in each representation's endeavor to become the sole object of consciousness.

Now, as they struggle to negate each other, each of the two opposing representations is, as Herbart puts it, inhibited (*gehemmt*) by the other (PsW: 130; SW V: 274, f.). We may liken his conception of consciousness, then, to a single, general psychic activity—call it "representing"—which, like the surface of a pool in a rainstorm, is disturbed by multitudinous sensory contact with realia. Each disturbance has the effect of differentiating the "surface"; and since that surface just *is* the activity of representing, each individual ripple takes part in that activity. Which is to say: instead of a general, undifferentiated representing activity of pure consciousness, we now have in each disturbance a particular instance of representing, i.e., *a representation*. Just as the ripples caused by two raindrops will interfere and conflict with each other, they nevertheless share a common substrate—the water—that makes both their being and conflict possible in the first place: they are ripples *of* water. So too each representation, as a ripple of consciousness, shares in the latter's essence, viz., the activity of representing (here the analogy to water breaks down, since water is not *per se* an activity). This active nature helps explain why Herbart often seems to suggest that individual representations possess their own energy or will[11] to survive and prevail over others. It is simply a manifestation of consciousness's own activity and "will" to maintain its essential simplicity, which is why he calls representations "self-preservings of the soul".

On Herbart's view, rather than thinking of *Vorstellungen* as static, image-like things or products of an antecedent act of representing,[12] we should instead think of them as individual conscious activities, i.e., atomic represent-*ings*.[13] Thus he can speak of the representation's "activity" persisting even when its "effect"—i.e., my awareness of it—"is held back by something alien to it". Now when an activity is held back or inhibited from expressing itself, we call it a "striving" (*Streben*). Therefore, since the many competing representations striving one against the others are all ultimately the self-preservings of *one* consciousness, *one* psyche, he groups them all together as single manifold within the same subject, as "one striving-to-represent [*ein Streben, vorzustellen*]" (PsW: 130; cf. LPs: 370; SW V: 274, f.). Herbart's theory seems to be confirmed by the common experience of our ideas brightening (*Erhellen*: SW V: 279; cf. Boudewijnse, et al. 1999: 170) and dimming (*Verdunkeln*, lit., "darkening", PsW: 130; cf. Boudewijnse, et al. 1999: 170), waning, vanishing, and suddenly returning. But experience does not reveal the *laws* according to which representations come and go in the mind. How can they be rendered calculable, and thus susceptible to scientific understanding?

Herbart proceeds as follows. As we have seen, insofar as representations conflict with each other in their struggle to find a place in consciousness, they inhibit (*hemmen*)[14] each other; an inhibited representation is correspondingly dimmed or obscured (*verdunkelt*); and to the degree that a representation is inhibited from full expression, it is transformed into a striving (*Streben*) (cf. esp. PsW: 133–4). Thus, it is reasonable, he argues, to think of the interplay of representations as psychological analogues of physical forces (*Kräfte*) (LPs: 369; cf. Boudewijnse, et al. 1999: 163; Weiss 1928: 74). When a representation dims, one can also say that it "falls" or "sinks" from awareness; when it brightens, that it "rises" into awareness (Weiss 1928: 75). It is important, however, to note

that he does not say that a representation really is a "force".[15] Rather, several representations *in opposition* behave in force-like ways.[16] A representation, then, expresses force only to the degree to which it is inhibited; and it is inhibited by another representation only if the latter is opposed to it in some way. Hence, just as this opposition does not belong essentially to either representation, so too their force-like behavior is only formal and contingent ("*Gradweise*", PsW: 134–5; cf. SW V: 278).

Representations oppose each other by degrees, as we see in the case of colors and tones (PsW: 135–6). This has important implications. Since the inhibitions immediately result from representations' mutual opposition, these inhibitions, too, are graduated, as are, consequently, the gradual dimming of the (represented) object and the transformation of a representation into a striving-to-represent: all these phenomena are to be thought of as graduated over a spectrum (PsW: 136). In short, the degree of opposition among representations sets one of two conditions for a degree of dimming. The second condition is the intensity with which each representation initially appears in the psyche. That is, representations

> may originally be either weaker or stronger, even without inhibition; we originally ascribe to all of our perceptions a certain degree [of native vivacity]. (SW V: 279)

In Herbart's conception of representations as opposing each other by degrees we find the cornerstone of his innovative application of mathematics to psychology, and thus of developing a "physics" of mind. He writes:

> If we now connect … the difference between representations in terms of their intensity [*Stärke*],[17] on the one hand, with the magnitude of their mutual opposition, on the other, then this yields in each case the magnitude of the resulting dimming, inhibition, [and] striving; and so too [the magnitude] of the remaining actual representation. Here calculation [*Rechnung*] finds its appropriate matter. (PsW: 136; cf. SW V: 279)

Extending his mathematical psychology by analogy to the theory of motion, Herbart complements his psychic statics (PsW: 139), or doctrine of representational equilibrium, with a psychic mechanics (PsW: 208), or doctrine of movement of these "mental bodies" over time (PsW: 137; cf. Boudewijnse, et al. 1999: 163). It is to these that we now turn.

3.3 Statics of spirit

As we have seen, on Herbart's view, representations are the perturbations, i.e., the individual or atomic representings that present as pictures the real objects that, through the senses, disturb the simple, continuous representing that is consciousness (PsW: 135). Any representation, taken in isolation, is *fully* determined by the real object it presents to consciousness. Thus, since "representation" just means the generation and maintenance of the presented picture, the concept of representation, as such, contains neither the notion of striving, nor of any activity directed beyond itself (PsW: 134–5). But this Parmenidean state of a single atomic presentation, a "well-rounded sphere" completely bathed in the light of awareness, never obtains in reality. Instead, various streams of new representations constantly come flowing into consciousness, and of these representations, many oppose each other. We might picture these atoms as elastic[18] balls streaming into the narrow beam[19] of awareness, jostling and crowding those already bobbing on the surface of consciousness. Depending on the strength and vivacity of these balls, some will remain floating, but most will quickly be pressed down, to some extent, through the surface and out of the light, which phenomenon was called "inhibition", above.

This picture, fanciful though it may be, can nonetheless help us get a grip on the idea of the "inhibition-sum" (*Hemmungssumme*), which plays a key role in Herbart's psycho-statics, but which he and his commentators treat in a very abstract, mathematical way. The idea is simply this: should two representations (to take the simplest case) oppose each other, then this opposition immediately causes them to squeeze and jostle each other, for they cannot both fully express themselves or be completely present to consciousness at the same time, as introspection readily confirms. Rather, they crowd (*drängen*) each other; indeed, each attempts to suppress or even completely crowd out *(verdrängen)* the other. As our picture suggests, this opposing and crowding is not a state but a movement, a "*Conflict*" or struggle to dominate, which takes time to reach a resolution (LPs, SW IV: 371 [§127]). The laws governing this movement are treated by mechanics (see following section), whereas statics deals with the end-state, i.e., the equilibrium that temporarily resolves the struggle, which Herbart terms the "static point" (*der statische Punct*) (cf. LPs, SW IV: 371 [§128]). This equilibrium is reached when each representation expresses itself to the greatest degree possible,

given the countervailing pressure of the other. The two elastic balls reach a state, in other words, where each presses or inhibits the other, but neither can increase that pressure further. So they come to rest in a state of mutual inhibition, in which neither "shines" as brightly as it would if unimpeded by the other.[20] What Herbart calls the "inhibition-sum", then, is the total of the opposed inhibitory forces, i.e., the total inhibiting exerted by the opposed representations upon each other, *in a state of resting equilibrium*. With this preliminary description in hand, let us turn to Herbart's mathematical presentation.

3.3.1 "Inhibition sum" and "inhibition ratio"

All psycho-static investigations are ultimately based on two determinations of psychic magnitude: (1) the "inhibition-sum" (*Summe der Hemmung* or *Hemmungssumme*); (2) the "inhibition relation", or better, "inhibition-ratio" (*Hemmungs-Verhältniß*).

3.3.1.1 "Inhibition sum"

In 3.3, I called the inhibition-sum the total of opposed inhibitory forces in the system of two (or more) opposed representations; Herbart himself likens it to

> the load [*Last*], generated by the oppositions among the representations, that is to be distributed [across the opposed representations]. (LPs: 371; cf. SW V: 284)

He uses a thought-experiment to derive the mathematical expression of the basic laws of representational "motion", that is, the laws according to which our mental representations interact with each other. Under normal conditions, representations are of different strengths, and so oppose each other with differential intensity, making a clear statement of their underlying laws difficult. Herbart therefore first considers an ideal case, the extreme at which two representations are "fully opposed" one to another, that is, one is completely inhibited, so that the other may remain unconstrained (PsW: 139). Under such circumstances,

> the most that can occur is the complete "sinking" of the former, i.e., a complete extinction of its represented content [*ihres Vorgestellten*] through the transformation of all its activity into a mere striving against its opposed [representation]. (PsW: 139; SW V: 281)

In other words, the former representation would be completely suppressed (*verdrängt*) (SW V: 281).

Herbart next posits a second idealized[21] case, in which the opposition is only partial, not "full", as above. In this case, he says, one representation could remain fully uninhibited even if the other were inhibited only to a certain degree (PsW: 139, ff.)[22] He now asks: Given representations *a* and *b*, by how much would each of the two completely opposed representations be inhibited? The inhibition of each individual representation may be found only if we *first* know the "*inhibition sum*" (*S*).[23]

> The sum of the inhibition [i.e., the total of the inhibitive force of both representations] is the quantum of the represent*ing* [*des Vorstellens*] that must be inhibited by the mutually opposing representations added together. (PsW: 140; SW V: 282)

Put another way, the inhibition sum *S* equals the inhibition of *a* (i.e., I$_a$��) plus the inhibition of *b* (i.e., I$_b$��), where both I$_a$�� and I$_b$�� are taken to be as small as possible.[24] (We should assume a minimal inhibition here since the natural state of representations is that of full liberty or radiance, which they always strive to reach as nearly as possible [SW V: 283].) Let us now return to the case in which representations, *a* and *b* are fully opposed. In this case, Herbart writes, "*either a, or b* must *be* the inhibition sum" (SW V: 282). If *a* is the stronger, then it will be *b* that is fully inhibited and dimmed. What is the minimal necessary magnitude of the inhibitory force that *a* exercises upon *b*? Exactly *b*, since beyond this amount there would be no cause for *a* to exert a dimming effect.

It turns out, however, that in either of the two cases (i.e., the case of full inhibition and the case of partial inhibition), the weaker representation turns out *not* to be totally inhibited, since a part of *S* must also fall onto the other representation (SW V: 282; cf. SW IV: 371, §130). Why? Because, after all, *S* is conceived as

> in effect the load distributed over the different representations that must bear it, viz., all of the antagonistically striving [representations]. (SW V: 282)

Hence, since a portion of *S* is borne by the countervailing representation, the weaker representation can (and must) remain uninhibited by just that amount. If we presuppose the total opposition of representations, then we can extend the above determination from two to any number of representations, so that *a* would be opposed by b,c,d,…n�,�,�,…�. *S* would then not be smaller than (b+c+d+…+n)(�+�+�+…+�), but also not larger.

> For if all those [representations] were completely suppressed, then the strongest [i.e., *a*] would remain completely uninhibited. (SW V: 284)

3.3.1.2 "Inhibition ratio"

Herbart next considers the proportion in which the inhibition sum is spread out over the various, mutually antagonistic representations; this proportion is what he calls the "inhibition-ratio" (SW V: 285, ff. [§43]). The calculation of the inhibition ratio may not consider an individual representation as "an originally aggressive force [*angreifende Kraft*], but only as a resisting force [*widerstehende Kraft*]", since, as we saw above, representations can on Herbart's view only be considered to be force-*like* insofar as they resist one another, not "in themselves" (SW V: 285). Nonetheless, a representation will better resist countervailing forces and assert itself against other representations, "the more vivacious [*stärker*]" it is (SW V: 285). Hence, he argues, the share of the load, *S*, that a representation must bear is inversely proportional to its vivacity: 1/i1/�, where "i�" signifies vivacity (SW V: 285–6; SW IV: 371).[25]

Earlier, we hypothesized *a* and *b* to stand in full opposition, where a>b�>� (SW V: 288, ff.). Hence S=b�=�, and *S* distributes itself over representations *a* and *b*. Let the part of *S* borne by *a* be called "α�", and the part borne by *b*, "β�".[26] Because b=S�=�, it follows that α� + β� = b. Since the inhibition sum *S* must for each representation stand in inverse proportion to its strength, it follows that α� : β� :: b : a (Weiss 1928: 77). These equations allow us to calculate the portions of *S* that are distributed to representations *a* and *b* individually.

$$\text{For a, the portion} = \alpha = \frac{b^2}{a+b} \quad \text{For } \Diamond, \text{ the portion} = \Diamond = \frac{\Diamond^2}{\Diamond+\Diamond}$$

$$\text{For b, the portion} = \beta = \frac{ab}{a+b} \quad \text{For } \Diamond, \text{ the portion} = \Diamond = \frac{\Diamond\Diamond}{\Diamond+\Diamond}$$

α� and β� represent the relative dimming of each representation (Weiss 1928: 77).

When we subtract the calculated portion of the inhibition sum from each original representation *a* and *b*, then the remainders "Ra��" and "Rb��"[27] designate

> the *degrees* of the remaining vivacity of the representations, after the inhibition has caused the previously calculated portion of actual representing to be canceled out [*aufgehoben*] and transformed into a mere striving. (SW V: 288)

Accordingly, what remains of the original vivacity following inhibition of *a* and *b*, respectively, can be expressed as follows (Weiss 1928: 78):

$$R_a = a - \frac{b^2}{a+b} \quad \Diamond\Diamond = \Diamond - \frac{\Diamond^2}{\Diamond+\Diamond}$$

$$R_b = b - \frac{ab}{a+b} \quad \Diamond\Diamond = \Diamond - \frac{\Diamond\Diamond}{\Diamond+\Diamond}$$

Now if the same conditions obtain for three representations, *a*, *b*, and *c*, where *a* is the strongest and *c* the weakest representation, then the inhibition sum S=b+c�=�+� (SW V: 288). As in the case of two representations above, we can calculate the remaining degrees of clarity (*Klarheitsgrade*) after inhibition (SW V: 288–9):

$$R_a = a - \frac{(bc)\cdot(b+c)}{bc+ac+ab} \quad \Diamond\Diamond = \Diamond - \frac{(\Diamond\Diamond)\cdot(\Diamond+\Diamond)}{\Diamond\Diamond+\Diamond\Diamond+\Diamond\Diamond}$$

$$R_b = b - \frac{(ac)\cdot(b+c)}{bc+ac+ab} \quad \Diamond\Diamond = \Diamond - \frac{(\Diamond\Diamond)\cdot(\Diamond+\Diamond)}{\Diamond\Diamond+\Diamond\Diamond+\Diamond\Diamond}$$

$$R_c = c - \frac{(ab)\cdot(b+c)}{bc+ac+ab} \quad \Diamond\Diamond = \Diamond - \frac{(\Diamond\Diamond)\cdot(\Diamond+\Diamond)}{\Diamond\Diamond+\Diamond\Diamond+\Diamond\Diamond}$$

3.3.2 *Limen* or "threshold of consciousness"

Herbart now calculates the remainders R� for different values for *a* and *b*, and for *a*, *b*, and *c*, showing that when only two representations, *a* and *b*, conflict, it is impossible for Rb≤0��≤0, i.e., for it to disappear completely from awareness (SW V: 289–90). However, according to Herbart's calculations, when three or more representations are in play, such that a>b>c�>�>�, it is possible for Rc�� … ≤0≤0 (SW V: 289–91).[28] This result, while perhaps obscure in its derivation, is the basis of Herbart's discovery of the concept of the *limen* or "threshold" of consciousness: for if, in

the opposition of representations *a*, *b*, and *c*, the inhibition sum *S* is distributed in such a way that the "remainder" of *c*'s vivacity is $\leq 0 \leq 0$, then this just means that *a* and *b* have completely dimmed or crowded out *c* from conscious awareness (cf. SW IV: 371, §130).

Still, he warns that we cannot attach any significance to a "negative representing", since the most that can befall a representation is total conversion into mere striving-to-represent, i.e., "that the remainder of actual representing $=0=0$".[29] Now if we let $R_c = 0 \lozenge \lozenge = 0$, then, substituting into the above mentioned equation (cf. SW V: 291),

$$0 = c - (ab) \cdot (b+c) bc + ac + ab 0 = \lozenge - (\lozenge\lozenge) \cdot (\lozenge + \lozenge) \lozenge\lozenge + \lozenge\lozenge + \lozenge\lozenge$$

$$c = bab + a - - - - - \sqrt{\lozenge} = \lozenge\lozenge\lozenge + \lozenge$$

This formula—the *Schwellenformel*—expresses what Herbart calls the "static threshold [*statische Schwelle*]",[30] at which representation *c* is completely eclipsed from consciousness by *a* and *b* (SW V: 293). Hence any further determination of *c* would be irrelevant for calculating the *S* for *a* and *b*:

> thus, the condition of consciousness, insofar as it may be statically determined, in no wise depends upon *c* … not to mention upon *even weaker* representations, which might be present in an infinite number without in the least being noticed in consciousness, so long as it was and remained in a state of equilibrium of all the representations. (SW V: 292)

Herbart immediately applies his discovery of the *limen* and the possibility of a representation's suppression beneath it. On the one hand, we notice, he says, that the mind is at any given moment preoccupied with a tiny fraction of the total number of cognitions, thoughts, or desires that *might* appear in our consciousness, given an appropriate stimulus (SW V: 292). This knowledge,

> absent, yet not annihilated, but rather remaining in our possession: in what condition is it within us? How does it happen that, although it is present, it nonetheless does not contribute to the determination of our state of mind until we recall it? What can prevent—sometimes for a long time—our most vivid convictions, best resolutions, and most refined feelings, from being effectual? What can imbue them with the unfortunate inertia by which they abandon us to vain regret? (SW V: 292)

The answer is simply: "*Other thoughts* have occupied us too vividly!" and therefore left that "knowledge" hovering beneath the *limen*. This simple solution, Herbart argues, in one stroke obviates the need for "false doctrines of transcendental freedom or of radical evil" (SW V: 292).

To sum up: One representation cannot so crowd out a second, since the remainder of b$\lozenge$, (i.e., [R_b] [$\lozenge\lozenge$]), can never $=0=0$. On the other hand, two representations suffice to crowd out a third completely out of consciousness, making it incapable of affecting the state of mind (*Gemüthszustand*); and this is all the more the case for further representations with a weaker vivacity than *c* (SW V: 292). The *limen* is determined as a limit ("*Gränze*") below which a representation is fully inhibited, but crossing which, it achieves a "degree of actual representing" (SW V: 292). The notion of a degree of representing allows Herbart to speak of representations "rising" and "sinking", and of representations lying more or less deeply beneath the *limen* (SW V: 293). "Consciousness", then, just *is* "the totality of coincident representing", and the representations at play in the arena of awareness make out the "state of mind" (SW V: 294; cf. SW IV: 372, §130). Thus Herbart concludes that the psyche's apparent property of being only able to activate a very small number of representations simultaneously *is no property at all*. Rather, it is the necessary consequence of the oppositions among our representations themselves, and thus dependent on their native vivacity (cf. Boudewijnse, et al. 1999: 178–9).

3.3.3 Fusions and complications

The soul's unity that we encountered in 3.2 functioned as the metaphysical principle of the inhibition of opposed representations in 3.3.1–2. Now, again, unity is the principle by which the unification of certain representations might be explained (cf. SW V: 307). We must keep in mind that our sensations (*Sinnesempfindungen*) form various continua, "sense dimensions" (Boudewijnse, et al. 1999: 179), or "modalities" (Boring 1950: 258), e.g., of color or taste and smell (PsW: 171; SW V: 307, f.). Within one and the same continuum, individual representations are opposed to each other,

while representations from generically different continua are not (SW V: 307).[31] Thus representations lying on one continuum will mutually dim each other to a degree (SW V: 308). So, whereas a number of colors, like "white, red, black, blue", will inhibit and dim each other, none of them can oppose and inhibit a taste representation (the example is at Weiss 1928: 80; cf. SW V: 312, f). Therefore, as representations may belong to the same or different continua, there must be, respectively, two different genera of representational unification (SW V: 308).

3.3.3.1 Fusions

In the former case, in which representations are of the same type, they will inhibit each other according to the degree of their difference and only join together so far as inhibition permits it (SW V: 308).[32] Their strength and opposition will determine "the law governing the precise nature of their unification", which Herbart calls the "fusion" (*Verschmelzung*) of representations (SW V: 308–9).[33] Examples might include blue and red fusing into the representation, "purple" (Boring 1950: 258). Two generically similar, and therefore opposed, representations will strive to fuse with one another, on account of the unity of the psyche. When they melt into one another as far as their mutual opposition permits, they reach what Herbart calls the "fusion point [*Verschmelzungspunkt*]" (Weiss 1928: 82).

3.3.3.2 Complications

Take two representations belonging to two distinct types, e.g., a color, a�, and a sound, α� (SW V: 308). Now, on the one hand, no inhibition is possible across continua, i.e., so a� will not dim α�. On the other hand, *ex hypothesi*, there is no second representation *within* each respective continuum to oppose and inhibit either of the two representations (cf. Felsch 1902: 5; Weiss 1928: 82). For these two reasons, they can unite *completely*, and are for purposes of calculation treated as a single force or single act, A� (SW V: 309; cf. Felsch 1902: 5).[34] Such unifications are "complete complications [*Komplikationen*]", or "foldings-together" (SW V: 308).[35]

Now imagine that the two heterogeneous representations a� and α� are joined by representations belonging, respectively, to the same types as those already present, e.g., color-representations, b� and c�, and sound representations, β� and γ� (SW V: 308). Here b� and c�'s appearance in the same color continuum as a� results in the mutual inhibition of a�, b�, and c�, so that b� and c� amount to what Herbart calls an "incidental obstacles" to the complete complication of a� and α� (SW V: 309).[36] (The same of course holds for α�, β�, and γ�.) Hence a� and α� can at best join in an "incomplete complication" (SW V: 309). Clearly these make up the vast majority of our compound representations (Felsch 1902: 5).

Finally: suppose that in one continuum, a� is inhibited by b� to a remainder r�; whereas in another continuum, α� is inhibited by β� down to remainder ρ� (SW V: 317).[37] Then

> remainders r� and ρ� will combine themselves to a total force [*Totalkraft*] which, however, cannot be separated from the [two] whole representations a� and α�, which are not completely connected. (SW V: 317)

If, a� were now threatened with further inhibition by a newly arising representation, c�, then because of the combined *Totalkraft*, α� would in effect "help" a� resist the force of c�.[38] Herbart calculates α�'s cooperative power at ρ·rα�·�� (SW V: 317–8; SW IV: 375).

3.4 Mechanics of spirit

3.4.1 Representational "motion"

Herbart's transition from psycho-statics to mechanics exhibits a counterintuitive order similar to that encountered above, in calculating the inhibition sum *before* calculating inhibition ratio (i.e., calculating the load before its distribution). Thus, we should note that psycho-statics describes the laws of a state reached *after* the laws of psycho-mechanics have operated on a given set of representations. To clarify: considered *per se*, representations are uninhibited, radiating at full strength (SW V: 338; PsW: 208). When two or more representations appear at once in consciousness, they conflict, and over time inhibit or dim each other's radiance until they reach equilibrium (SW V: 338).[39] That balanced *end-state*, and the distribution of the inhibitory load upon each representation at that end-state, is the concern of psycho-statics. Psycho-mechanics, by contrast, studies the laws of representational motion, i.e., the laws governing the preceding conflict between representations that *leads to* equilibrium (SW IV: 371). By "motion", Herbart means the falling (and rising) of representations with respect to the static threshold, or, more precisely, their dimming (and brightening) with respect to the threshold of awareness (SW V: 338, 339; Boudewijnse, et al. 1999: 181, ff.; Weiss 1928: 84. Cf. esp. Felsch 1904: 194, f.).

This initial précis of the relation of psycho-statics to psycho-mechanics is helpful for understanding Herbart's claim that "the inhibition sum [*S*] must, in accordance with its law, continuously fall" (SW IV: 341; cf. SW V: 338; Felsch 1904: 195; Boudewijnse, et al. 1999: 182; Weiss 1928: 83, f.). This statement, on the face of it, is deeply puzzling, and points to a vexing ambiguity in his usage of the term "inhibition sum". For by speaking of the "falling" (i.e., diminution) of the inhibition sum, Herbart gives the impression that it is variable. But as our discussion of the inhibition sum above

showed, Herbart in psycho-statics conceives it as the sum of the (amount of) inhibition each of a number of conflicting representations must partially "bear" for them all to reach a state of equilibrium. But this sum is in fact a fixed quantum,[40] a constant, dictated by the inherent vivacity of each representation and the degree of opposition[41] among representations. In a case examined above, S is the load *actually borne* by a and b when they reach equilibrium (cf. 3.3.1.1).
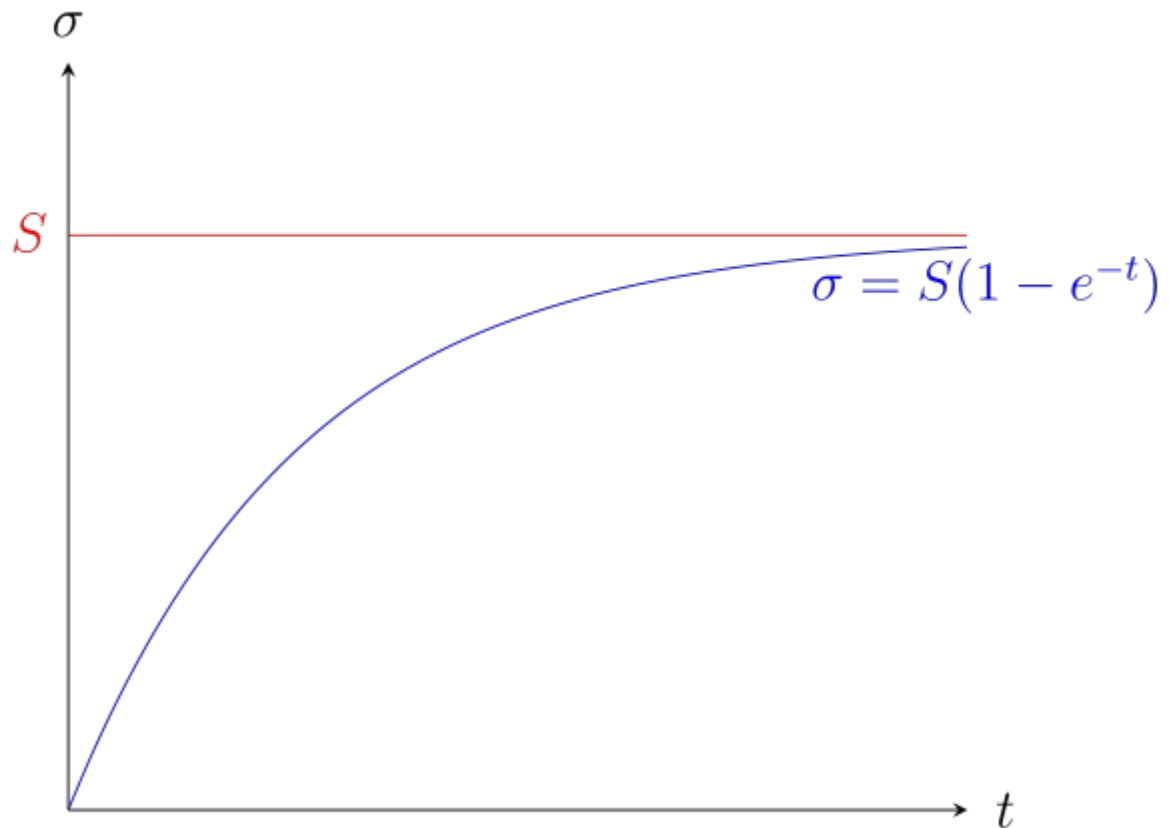
But psycho-mechanics is concerned not with the condition of a and b *in* equilibrium, but *on the way to* equilibrium. What phenomenon can we witness here that *could* be called (loosely) the "falling of the inhibition sum"? Consider now not the end-state of equilibrium, but rather the beginning of the process. Here, at t₀�0, one representation a is joined by a second representation b. At this first instant, the vivacity of each representation is at its uninhibited maximum,[42] and therefore the *forces* of opposition are also maximally active. Notice now that while, as stated above, S, the load of forces distribut*ed* at equilibrium is a constant already fixed by the nature of a and b, these forces, at t₀�0, are not at all distributed but active, with a and b bearing *no load* (as yet). Hence, at t₀�0, one could reasonably call these active forces the "inhibition sum", but not in the sense of the distribut*ed* load, but rather in the sense of the load-*to-be-distributed* over a and b. Let us call the distributed load "S", to remain consistent with the use of the term in psycho-statics, but let us call the load-to-be-distributed "Sₙ��", to indicate the load to be distributed at tₙ��.[43]

As we saw above, Herbart argues that from the instant at which a and b "collide" in consciousness, to the point of static equilibrium, a period of time must elapse. While at t₀�0, the oppositional forces are at a maximum,[44] at t₁�1 they will already have dimmed each other to an extent; at t₂�2 this dimming will have progressed even further; etc. Herbart symbolizes the *combined* amount of the dimmed portion of a and b with "σ�".[45] Hence we can see that as time passes and a and b progressively inhibit each other, Sₙ��, the load-to-be-distributed, continuously diminishes: this is what Herbart means by saying that the inhibition sum falls. And as Sₙ�� falls, σ� increases, that is, the inhibited amount of a and b increases since they are taking on as a load a portion of Sₙ��. He notes, moreover, that as Sₙ�� falls—i.e., the forces' conflicting activity decreases—the inhibitory pressure they exert also decreases (SW V: 339, e.g.). Hence, the *rate* of inhibition and dimming decreases over time as well. The situation can be summarized thus: at first a and b stand in maximal conflict, leading to rapid absorption of these forces in the form of dimming; hence, as the active forces decrease, the rate of dimming also decreases, as a and b jostle ever more slowly into equilibrium around the static point (cf. Felsch 1904: 196). As Felsch writes, "the billowing sea is becalmed but by and by" (Felsch 1904: 195).

Now Herbart, in his LPs, rightly calls the laws of psychic motion "highly diverse and for the most part very knotty", and restricts his treatment to just one equation:[46]

$$\sigma = S(1 - e^{-t})\mathbf{\square} = \mathbf{\square}(1 - \mathbf{\square} - \mathbf{\square})$$

where S is the inhibition sum (as eventually distributed; ≠ "Sₙ��"); σ� is the combined amount of a and b that has at any time t been inhibited; and e is the base of the natural logarithm.

We can see from the graph that as t grows towards infinity, e^{-t} diminishes towards zero, and so σ approaches, but never reaches S. That is, the total amount of a and b that is *actually* inhibited never in a finite time attains the inhibition sum; which, in turn, means that a and b never reach complete equilibrium (SW IV: 372).

> Due to this latter circumstance, in a waking person, even in a state of equanimity, the representations are always caught up in a gentle floating. (SW IV: 372)

Now one might think that as the representations jostle towards equilibrium they will become less "conscious [*bewusst*]" (i.e., less perceptible to consciousness), and of course this is true for each inhibited representation with respect to its own (hypothetical) uninhibited start-state. But at issue here is the clarity of the representations *joined in opposition*. A moment's reflection suggests that at the moment of collision, when the inhibition sum S_n is highest, neither representation will be at all clear to consciousness, but on the contrary, be least perspicuous; considered *separatim*, the representations will be obscure. Therefore, as S_n falls over time, the representations gradually "rise", becoming more and more perspicuous and clear, though less *vivacious* than in a solitary state (cf. Felsch 1904: 195; Weiss 1928: 83).

3.4.2 Reproduction of representations

As we saw earlier, Herbart considers representations to be the "self-preservings" of the psyche expressed in its, the psyche's, own nature; their multiplicity stems from those disturbances which the psyche resists at each moment of its self-preservation (SW V: 387). Representations, once formed, must somehow remain in the psyche, or else no self-consciousness would be possible (SW V: 354; 387). Whenever a given disturbance lasts for a certain time, the new representations that are constantly springing into being at every moment begin to accumulate (SW V: 387). For the psyche's self-preservings and the representations are one and the same, just in different relations (SW V: 401):

> By the word, "representation", we … mean the phenomenon insofar as we may encounter it in consciousness; by contrast, the expression, "self-preservation of the psyche [*Seele*]" refers to the real act [*realer Actus*], which immediately brings forth the phenomenon. This real act is not an object of consciousness, for it is the activity itself that makes

consciousness possible. Thus "self-preservation of the psyche" and "representation" go together as action and occurrence [*Thun und Geschehen*; i.e., action (and simultaneously resulting) occurrence]. (SW V: 401)

The "soul's" self-preservation also continues when a representation has for a certain time been forced out and supplanted (*verdrängt*), which is why it can return to consciousness once the occluding factor goes away (SW IV: 376). This is the basis of those mental phenomena that Herbart calls "reproduction of representations" (SW IV: 376).

The conditions under which a representation reappears in consciousness determine whether its return is called an "immediate" or a "mediate" reproduction (cf. Boudewijnse, et al. 1999: 181; 185, ff.). An immediate reproduction occurs when a hitherto inhibited representation strives back up into consciousness under its own power, as it were, as soon as the occluding factor has been neutralized (SW IV: 376). Herbart writes:

It is usually the case that a new perception allows the older representation of the same or very similar object to step back into the light. This occurs in that the new perception pushes back everything that happens to be present in consciousness currently opposing the older representation. Immediately the older [representation] arises of its own accord without further ado. (SW IV: 376)

In the case of mediate reproduction, on the other hand, a representation reenters consciousness not from its own power, but through the aid of representations with which it is somehow connected. The partial representations mutually help each other as soon as, and to the extent that the pressure subsides that suppressed the connection as a whole (SW IV: 377). Taken together, these two types of reproduction furnish the foundation for recollection and memory.[47]

3.4.3 Representational series

Mediate reproduction or recollection plays an important role, moreover, in the formation of representational series or sequences (*Vorstellungsreihen*). Time and the quality of the representations determine the formation of such series (SW V: 410). Assume that perceptual representations a, b, c … appear sequentially in consciousness, and are not mutually opposed (SW V: 410). Representation a, which enters consciousness first, is quickly opposed by other representations, which inhibit and dim it, causing it to sink down to the degree ("*Klarheitsgrad*", Weiss 1928: 86), r. Next, representation b joins it and fuses with a's remainder, r◆'. Together, they both continue to sink, with b now becoming R and r◆' becoming r″◆″. Representations c and d … join them. The following schema becomes apparent (SW V: 410):

arrʳr‴…r(n)bRRʳR″…R(n−1)cρρ′…ρ(n−2)dr…r(n−3)etc.etc.◆◆◆′◆◆◆″◆′◆◆◆″◆″◆
′◆etc.…………◆(◆)◆(◆−1)◆(◆−2)◆(◆−3)etc.

Now suppose the entire series were suppressed to the *limen*. If *one* representation-member of the series were allowed to rise back into consciousness, it would have a "reproductive effect" upon all its fellows (namely in accordance with the laws of reproduction laid out in §88; SW V: 372, ff., esp. 375, f.). If the first representation, a, rises as an "immediate reproduction", it then most quickly pulls b up after it, since b is most completely fused with r; but somewhat more slowly c, and with ever decreasing speed, d, etc. (Stout calls this the "evolution" of a series; Stout 1888a: 335). But if c, that is, a representation somewhere in the middle of the sequence were the first to resurface, then, because it is *fused* with r◆' and R—the remnants of a and b—these latter will be dragged past the *limen* simultaneously (Stout calls this the "involution" of a series; Stout 1888a: 335); on the other hand, d, e, … will follow in succession (SW V: 410–11).[48]

3.5 *Geist* and *Gemüth*: the life of the mind

Herbart is of course very concerned to demonstrate the superiority of his system over other, traditional conceptions of mind (SW VI: 53). In particular, he wants to head off critics of mathematical psychology who argue that "mathematics merely determines quanta, whereas psychology deals primarily with qualities" (LPs [1850]: 29). He therefore turns from the statics and mechanics of *Geist* (soul as representing) to the special theory of the "*Gemüt*" (soul as feeling and desiring), in order to show how the latter can be incorporated into the former (LPs [1850]: 29; SW VI: 56, f.).

For Herbart, *Gemüt*—which I will here translate as "disposition"—has its seat within *Geist*; put another way, it is our *state of mind*. Hence,

> feeling and desire are in the first place states
> [*Zustände*] of representations, to wit, for the most
> part, *mutable* states thereof (LPs [1850]: 29)

when we feel, something is always simultaneously represented, and the feeling is contained *in* that representation (LPs [1850]: 32).[49] Again, when we desire, we always also represent that which we desire (cf. Aristotle, *Metaphysics* Λ7, 1072a26, ff.).

Herbart develops his theory of disposition on the basis of persisting (*stehende*) and rising (*steigende*) representations (SW VI: 57).[50] That is, all dispositions must ultimately be explicable in terms of persisting or rising representations, and nothing else. Taking first the case of a representation (or representational series) *R* that stands still in consciousness, its persistence may be due to the simultaneous and countervailing forces exerted on it by rising and inhibiting representations, Rr�� and Ri��, respectively. Hence we will *feel* the state of *R* as a "compression [*Klemmung*]" between Rr�� and Ri�� (SW VI: 64; 58; cf. LPs (1850): 31) and such states are "unpleasant feelings" (SW VI: 59). On the other hand, if a representation *R* rises, then it may encounter an obstacle that is not strong enough to prevent it from rising, or it may be aided by advantageous forces (SW VI: 58–9). Since these forces affect the real activity of representing, Herbart argues that we cannot be unconscious of their pressure upon or aid of *R*:

> But these [impingements upon *R*] are not
> [themselves] *objects* of representing, but merely
> the *way* in which the representing is happening; these
> determinations of consciousness, insofar as it
> transcends the mere representing as such, must be
> called "feelings" [*Gefühle*]. (SW VI: 58)

What then, on Herbart's view, is desire? "Feelings" accompany representations that, as they rise, become continuously more "effective" (SW VI: 59). By "effective", he means that the representation in question determines more and more representations akin to it, while at the same time inhibiting those that are not (SW VI: 59). "Desire", then, is the consciousness of a representation's rising struggle towards ever-brighter light, and of its corresponding "tension" as it works itself up and over obstacles (SW VI: 59).

Herbart extends his interpretation of feeling and desire to the phenomena of affect[51] and passion, the "strongest expressions of feeling and desire [respectively]" (SW VI: 75). In both cases it is important to see that the *force* of both affect and passion inheres in the representations themselves; when these forces are compressed or relaxed beyond the static point, they produce tension within the masses of representations, which manifests itself as affect (Weiss 1928: 90). Both result from movements, not of individual representations, but of fused representational masses well past their point of equilibrium (SW VI: 75; cf. Weiss 1928: 89). Thus, if a "larger amount of actual representation is raised into consciousness than can subsist there [all at once]", the result is an "ebullient" affect (SW VI: 75).[52] By contrast, if a larger mass is banished beneath the *limen*, the psychic barometer falls as well, resulting in a "melting" affect (SW VI: 75). Just as low atmospheric pressure provokes stormy weather, so too the disappearance of a representational mass loosens and unleashes those masses that remain (SW VI: 75–6). These suddenly far exceed the static point, their violent expression taking the form, e.g., of rage or fear (SW VI: 75–6). This also explains why affects are always transient moods, since representations always strive to return over time to a state of equilibrium (SW VI: 75).[53]

3.6 Spatio(-temporal) representations

Because our desires and feelings are inextricably intertwined with representations of our environment, we are naturally led to inquire into our perception of things in the world; this in turn requires an analysis of *types* of spatial and temporal representation (SW VI: 86). Not concerning himself here with the abstract concepts of space and time, Herbart instead looks to the genesis of spatial perception (*räumliche Auffassungen*) in earliest childhood. He notes that children irreversibly acquire, through practice, such perception:

> The child's hand first *learns* to grasp, the eye
> first *learns* to direct itself appropriately; but the
> adult involuntarily achieves what he learned;
> moreover, the *pure sensible perception* [of the
> child] is obscured by additions that his training
> [*Ausbildung*] mixes in. (SW VI: 86)[54]

Similarly, although there are those who can do without sophisticated, acquired concepts of temporal magnitudes, roughly noting the slower and the faster, yet once we learn rhythm and the division of time, we cannot return to the more primitive state. Herbart thus disputes the notion that space and

time are originary facilities or faculties: the spatial and temporal are rather *acquired*, "trained, even artificial conceptions" (SW VI: 86).

Herbart now faces a problem. All spatial and temporal differences, such as right and left, or earlier and later, lie in the representation, not in the act of representing (SW VI: 87). Yet because represent*ing* must precede the represent*ed*, the soul must, through its representational activity—which, itself, exhibits no spatio-temporal differences—introduce spatio-temporal relations into the representation. Yet the representation is not real; and, *quā repraesentātum*, it cannot "really [*wirklich*] be separated out in space or time": "the actual [*wirkliche*] psychological event of spatial representing is something altogether non-spatial", just as "the representing of the temporal is something in which there is nothing of the *thereby represented* time" (SW VI: 89).[55]

Herbart addresses the problem by appeal to the laws of reproduction (treated in the section on psycho-mechanics, above) (SW VI: 89). The represent*ing* (activity) of the spatial must in *some* way resemble the spatial itself, or else its spatial *repraesentātum* would in no way be "spatial". And this resemblance lies in the multi-dimensional nature of the interweaving of representational series. As Stout simplifies Herbart's presentation,

> interweaving means that from each term of a given series[,] trains of reproduction start, which are in their turn at once separated and interconnected by cross-series. On the side of the presented content[,] regular mechanical interweaving is correlated with the consciousness of a spatial order.... (Stout 1888a: 338; SW VI: 89)

It is in this interconnection of "cross-series" that Herbart identifies the appearance of a lateral or "neighboring" ordering (*eines **Neben einander** geordneten*). The crucial element in this order (which applies not only to space but also to time and number (cf. Natorp 1910, and esp. SW VI: 110, ff. (§116)) is the "gradation in the representations' connection [*Abstufung in der Verbindung der Vorstellungen*]" laid out in psycho-mechanics (above) (SW VI: 90). Only now, as we apply that abstract doctrine to experience, we see that the small number of representations used for illustrative purposes (a,b,c,d,...)(�,�,�,�,...), together with their dimmed remainders (r,R,ρ,r...)(�,�,�,�...), are far too few to account for the innumerable representations generated by the *minima visibilia* and *tangibilia* of actual experience (SW VI: 90). Herbart writes:

> With every sense-contact of the spatial, every little colored or touchable place [*Stelle*] produces its own representation; and every representation fuses with every other. (SW VI: 90)

The picture he paints here is astonishing: as we see and touch the world—in particular, as one "moves the observing eye and the touching finger *forwards* and *backwards*"—not only do we generate an unfathomable manifold of visible and tangible representations at any time *t*, but each of these representations in turn produces a train of innumerable fading, graded fusions with all the others (SW VI: 90). Thus innumerable series are generated and become interwoven (Weiss 1928: 92). The constantly generated perceptual representations gradually sink and shade away, fusing less and less with those trailing after (SW VI: 90). But at the slightest return (i.e., due to the finger or eye moving in reverse of its initial direction), all the earlier percepts, now aided by the very similar new-comers, begin to rise, which rising is associated with

> a *nisus* [upward pressure] to reproduce all those remaining [representation], whose [rising] velocity has exactly the same gradations as the earlier [graded] fusions. (SW VI: 90)

How does this picture begin to explain the spatial character of visual and tangible representations? The key is that *any* of the earlier acquired percepts can, in rising, pull up over the *limen all* the others, to wit, in their ordered places next to or between each other, and this all the while that the *act* of representing remains "purely intensive" (SW VI: 91). Thus the spatiality of our (visual) representations is a function not of immediate sensation, but of an interior process that lends perception "spatial form", namely those already present representations that reappear, governed by a law according to which each representation affects, in a regular way, the appearance of the others connected to it (SW VI: 91). Herbart writes:

> Insofar as the instantaneous [new] sense-perception [*Wahrnehmung*] fuses with the already ordered representation, it too becomes ordered; and in this way

> the persisting perception is incorporated in a
> continuous transition to spatial form. (SW VI: 91)

Temporal representations are generated in a similar way, viz., in accordance with the laws governing the reproduction of the continuously woven tissue of representations (for a more detailed account of the generation of temporal representation, see Stout 1888b: 474, ff.).

3.7 Higher cognition

3.7.1 Thinking things

We do not perceive "the spatial" or "the temporal" as such, but only ever *things* in space, or *events* in time (SW VI: 112).[56] Herbart again explains this fact in terms of the "complete and sufficient explanatory ground" of the psyche's unity (SW VI: 116). The soul's unity necessitates that our representations of things, and the events taking place among them, originally constitute themselves out of the simple sensations of the particular senses. Therefore, all our representations would present (*darstellen*) only a single object that would be neither spatial nor temporal, without parts, and immune to differentiation—except for the inhibitions and oppositions that underlie the simple sensations of the particular senses in their mutual encounters in the psyche (SW VI: 115, e.g.; cf. Stout 1888b: 478). This is what conditions the multiplicity of things and events: they are nothing other than the determinate groupings of sensations *in our consciousness* (Weiss 1928: 93).

3.7.2 Conceptual thinking

But if our sensible representations would, in themselves, present a simple, partless, Parmenidean object beyond time and space, how is it that they in fact seem to represent distinct *things* in our awareness? At first glance Herbart would seem to be presenting the Kantian picture of a chaotic sensible manifold, upon which a spontaneous Understanding could work its synthetic operations. But in keeping with his anti-faculty position, Herbart argues instead for a kind of "passive synthesis" *avant la lettre* (SW VI: 116), by which certain representations are connected and collected into separate unities, but spontaneously, motivated solely from within, in accordance with the laws of psycho-mechanics and psycho-statics (cf. esp. SW VI: 114, ff.).

Herbart proceeds by giving a "naturalized"[57] account of concept formation, that is, of *representation-masses* that, over time, develop into instruments by which we become conscious of *things*, and ultimately of ourselves. Thus, while there is "no doubt that just as concepts arise out of sense-perceptions [*Wahrnehmungen*], so, too, do clear concepts arise out of unclear[58] concepts"— nevertheless, these processes must be explained without recourse to a *deus ex machina* one calls "the Understanding" (SW VI: 117).

Anticipating the battles over logical psychologism of the late 19th and early 20th centuries, Herbart carefully distinguishes between a psychological and a logical sense of "concept" (SW VI: 119–20; cf. Stout 1888b: 477; van der Schaar 2013: 83). He writes:

> Every thought [*cōgitātum*], considered merely
> according to its quality [sc. *quā cōgitātum*], is a concept
> in the logical sense. (SW VI: 119)

By "concept in the logical sense", he means the represented

> content considered apart from the psychological
> conditions and circumstances of its presentation at this
> or that time to this or that individual mind. (Stout
> 1888b: 477)

Thus, concepts offer "a common knowledge for all men and all times", and, as such, are *nothing psychological* (SW VI: 120). Psychology, by contrast, investigates *real* thinking, and tries to determine how concepts come to be *actually* thought. As a thought-content of consciousness, then, a "concept" is a representation, namely one "which has as its represent*ed* [content] the concept in the *logical* sense" (SW VI: 120), or, as Weiss puts it, "that representation through which the *repraesentandum*[59] (i.e., the logical concept) is represented in reality" (Weiss 1928: 94).[60]

What, then, distinguishes a (psychological) concept from other mental contents, like sensations, imaginings, or memories? Because the soul is, *ex hypothesi*, a *tabula rasa*, there are neither original concepts (like Kant's categories) nor faculties of concepts (like Kant's Understanding or Reason): "*all concepts are things that have come to be*" (SW VI: 120; cf. esp. SW VI: 129, ff.). The incipient concept must therefore find its origin in the psyche's attempt to preserve itself, that is, its inherent simplicity, in the face of an outside disturbance. A representation in the process of coming to be is called a sensation (*Empfindung*) or perception (*Wahrnehmung*) (SW VI: 120). Upon entering consciousness, it is immediately subject to inhibition by other, already present representations. When, under certain circumstances, it returns to consciousness, the representation is called an "imagining [*Einbildung*]", and, as we saw in our discussion of representational reproduction, when this reproduced imagining connects with a series of fused (temporal) representations, it may become a memory (SW VI: 121).

When, then, do concepts arise? They are not generated as individual representations at a particular point in time, nor do we have them in addition to sensations and imaginings. Rather,

> we ascribe concepts to ourselves insofar as we abstract from the entrance of our representations into consciousness, and instead reflect upon the fact that they exist there, and now in fact allow their *repraesentātum*[61] (the concept in the logical sense) to appear [in consciousness]. (SW VI: 121)

So, if the psychological concept is to have as its *repraesentātum* just the content of the logical concept (its *repraesentandum*) and nothing more, then the psychological concept would have to shed all its peculiar and contingent traits, all of the complexions and fusions that have inevitably accrued to it in its own generation and reproduction (SW VI: 121). How is this possible? First, through a process of isolation into a "crude" or "confused" concept (SW VI: 125); second, through the crude concept's analysis by means of judgments; and third, through the classification of these judgments (cf. Stout 1888b: 477; Weiss (1928) overlooks the first step).

Isolation occurs as many representational series are progressively curtailed, in accordance with the laws of reproduction. Thus, if there we take two similar sensations, *a* and *b*, Herbart argues that as *b* is occurring, *a* reproduces itself as an imagining, together with all its fusions and complications. [62] If now a third similar sensation, *c*, appears, then *a* and *b*, together with their connections (and connections of connections) reproduce themselves, but "now there is already an inhibition, as the connections of the one and the other will differ" (SW VI: 123). Imagine now that new, similar sensations flow in by the hundreds or thousands: "it is evident that then diverse associations of all the preceding [sensations] will be as good as extinguished in their reproduction", letting only a tiny amount of each enter consciousness (SW VI: 123). Nevertheless, all these small amounts add up to a considerable total vivacity (*Totalkraft*), and this fused representation is tantamount to a concept. In other words, as similar elements of representational sequences fuse, the opposing links of these series cancel each other out, leaving only the fused total force shining forth surrounded by a "dim margin of competing alternatives" (Stout 1888b: 478). Having seen one and the same man in various positions, moods, clothes and places, then, should we see him again, the *total*-representation that now appears is the concept of this man.[63] According to Herbart, crude general concepts of objects and events arise analogously to these incipient individual ones (SW VI: 124–5).

Having shown how incipient concepts arise as the strong, common nodes of reproduced series of similar representations, he proceeds to show how, given such nodular representations, they can be definitely discriminated into subject and predicate, viz. a judgment (Stout 1888b: 478). He observes that the assimilation of a new sensation to earlier, similar representations can be regarded as (analogous to) a judgment of a crude psychological, if not of a logical sort:

> the sensation provides the subject; the fusion is … the copula; the earlier representation now fusing with [the presently occurring] sensation takes the position of the predicate. (SW VI: 126; cf. Stout 1888b: 479)

Although in this case, the subject and predicate fuse so quickly and imperceptibly as to be indistinguishable, yet it provides the template of the *conscious* separation of subject and predicate, and hence of logical judgments proper. To wit: if and when the fusion is somehow hampered or delayed, then its beginning, middle, and end are able to "have their say" (SW VI: 126). Specifically, at first the newly arrived representation (i.e., the sensation) of the subject subsists alone in consciousness. As it starts to sink, the representation of the predicate rises, exerting an influence—partly inhibiting, partly promoting—on the representational sequences extending out from the subject as they strive to evolve. "The copula appears in the middle as the expression of that change in disposition which takes place in the fusion" of the subject and predicate (SW VI: 126).[64]

3.8 Apperception

3.8.1 Background

Herbart is perhaps most famous for his theory of apperception and its application to the problems of the ego and self-consciousness (Boring 1950: 256–7; Stout 1888b: 498). He considers these notions and their metaphysical elaboration as faculties or powers to be nonsense, giving rise to paradoxes of self-consciousness; for example, "a faculty of self-observation [introspection]" raises the specter of an infinite regress of "self"-observing powers (SW VI: 140). This is not to deny that ego, self-consciousness, or self-perception answer to genuine psychological phenomena. Rather, it is to set a task for psychology, namely how to account for these psychological facts solely in terms of representations and without recourse to faculties, since

in the soul there are only representations; and out of these everything must be composed that is to appear in consciousness. (SW VI: 141)

Because Herbart's theory of apperception is ultimately about the formation (*Bildung*) or structuring of the ego, it also provides the cornerstone of his theory of education (*Bildung*) (Hayward 1907: 15; Boring 1950: 257; Weiss 1928: 98).

As we have seen, representations never appear singly, but continuously interact, complicating and fusing with or inhibiting each other. They form sequences and series, which, in turn, get interwoven with each other, forming stable groupings called representational masses. Simply put, "apperception" is Herbart's term for the assimilation of one representation or representational mass into another. In this section I will explain not only the mechanism of this process, but also why Herbart calls it "apperception".[65]

3.8.2 Mechanics of apperception

We earlier dealt with the fusion and complication of representations, i.e., with the laws governing the formation of representational series and masses. We are therefore already familiar with the laws of the assimilative process called "apperception", but only from the "outside", observing the representational masses objectively. To understand the assimilative process as "apperception" proper, we must now recall that representations are not (merely) represent*ed* "objects" of consciousness (*repraesentāta*), but (also)—as perturbations of the soul—represent*ing* "subjects" of consciousness (*repraesentantia*). In other words, representations, as *bewußt* or conscious,[66] must be (also) conceived as "observing", of "being-aware" (*beobachtend*). Always keeping this peculiar characteristic of Herbartian representations in mind, let us consider his interpretation of the phenomenon that has been (in his view) mistakenly called "inner sense" or a "faculty of self-observation [*Selbst-Beobachtung*, i.e., introspection]". He writes:

One representation (or representation-mass) is observed; *another* representation (or representation-mass) is the observer. (SW VI: 141; emphasis added)

In order to support this claim, Herbart considers two cases: (a) the assimilation of a sense-representation (a "sensum"), called "outer apperception"; and (b) the assimilation of a reproduced (i.e., recollected) representation from below the *limen*, called "inner apperception" (Stout 1888b: 448). The latter is treated as analogous to the former (Stout 1889: 448).

3.8.2.1 Outer apperception

A new sense perception or perceptual series, W,w,w,ω�,�,�,� ("W�" for *Wahrnehmung*, sense perception), always appears in an arena of consciousness already occupied by some older representation or representational group, R,r,r,ρ�,�,�,� (SW VI: 143). At first, the "unarrested intensity" of W,w,…�,�,… gives it as it were the drop on the pre-existing representations, initially forcing them down against the threshold (Stout 1888b: 484; SW VI: 143). At the same time, in accordance with the psycho-mechanical laws of reproduction, elements of the new series (e.g., w,w�,�) reawaken the subliminal *similars* amongst the older representations (e.g., r,r�,�), and fuse with these, while simultaneously inhibiting the *dissimilars* (e.g., R,ρ�,�). But the similars (r,r�,�), being themselves more strongly bound to the dissimilars (R,ρ�,�) (as members of the same mass), gradually draw up the latter, and so the whole mass (R,r,r,ρ)(�,�,�,�) returns more strongly to inhibit W,w,…�,�,… (cf. esp. Stout 1888b: 484). As Stout points out, the returning pre-formation (R,r,…)(�,�,…) would seem likely to suppress W,w,…�,�,… except for their "points of affinity":

So far as they are homogeneous they fuse. The appercipient group retains in consciousness what is kindred to itself, and at the same time represses what is antagonistic. (Stout 1888b: 485)

Hence

the sense-given group [W,w,…][�,�,…] becomes [after considerable modification] incorporated with the pre-existing system of presentations [R,r,…][�,�,…]. (Stout 1888b: 485)

3.8.2.2 Inner apperception

The process of outer apperception provides the model for inner apperception (SW VI: 141). Simply replace the *sensa* W,w,…�,�,… with a relatively weak representational series m,n,o,p,q,… �,�,�,�,�,… already present in (as opposed to newly entering) the soul. If now there is in

consciousness at the same time a second "stable and powerful mass", P,P,p,π,...�,P,�,�,..., the two series will complicate with or "fold into" each other, generating some new complex (such as Pm�� or Pn��) that has elements from each series (Stout 1888b: 485). Now from each member of the new complex, two chains of reproduction will evolve, e.g., in the case of Pm��, one from P�, and another from m�. Each of these two chains strives to unfold itself according to its own law. But because both series contain the common member p�, the evolving chains must fuse at this node and so reinforce each other. As in the case of outer apperception, the *m*-series is initially more vivacious and depresses the *F*-series. As they interact, the greater strength and coherence of the *F*-series reacts, "[re-]form[ing] the [weaker *m*-series] in its own image": the stronger mass preserves the similar elements in the weaker by fusing with it, while driving it back at other, dissimilar points (SW VI: 144). From the point at which the two chains are fused at p�, their respective reproductions cannot continue, as they would have separately. Instead, they now become interwoven into a new representational mass with "a new total-force" (SW VI: 141).

In both inner and outer apperception, we find a stronger and more stable structure into which a relatively unsettled and hence weaker element is integrated. They are drawn together at nodes of similarity, while dissimilarity blocks a complete fusion. As a result of this contest of forces, the weaker element gets stuck or "fixed" by the stronger. Now recall that both stronger and weaker elements on this account are *representings*; therefore the stronger one will be the observ-*ing* ("awar-ing") consciousness, while the weaker one, while not *un*-conscious (for it is stuck *above* the *limen*), is the observ-*ed*. In the above example, the *m*-series will be "observed" "like an object" by the *F*-series (SW VI: 144).[67] It is just this unbalanced relation between representations (or representational structures and elements) that constitutes the phenomenon of self-perception, i.e., apperception.

3.8.3 Apperception and perception

Herbart's account of apperception is closely linked to the phenomenon of attention, for, as Stout puts it, "to be in consciousness is not the same thing as to be an [explicit] object of consciousness":

> [t]here is always a dim margin of [re]presentations, which, though not distinctly attended to, are none the less components of the total conscious state. (Stout 1888b: 485)

Consider the following situation: one is engrossed in an activity like reading or a game, and so fails to notice a background noise, like a refrigerator's hum. Only when the hum "suddenly" stops do we realize that it had been "in our consciousness" all along, yet without being an "object of consciousness". Hence, if to perceive is to attend, then sense perception depends on an apperceiving mass (Stout 1888b: 486). It is only when the dimmed representation begins to fuse with the vivacious grouping at the center of awareness that the former

> acquire[s] that peculiar distinctness, which we express by saying that we notice it, observe it, or attend to it, or that it is an object of notice, observation or attention. (Stout 1888b: 485)

Thus, in one respect, sensation always precedes apperception, namely insofar as all representations and their groupings originate in sensory disturbances (SW VI: 143). Yet genuine *perception*, even of *sensa*, depends in turn on apperception, for only apperception allows us to *attend to* the sensed "as" or "in terms of" the apperceiving mass to which it is being assimilated (Stout 1888b: 485, f., §29). As we saw in Section 3.7.2, this is just to say that we need a representational mass called a "concept" in order to recognize a new representation as an instance of it, and thus in the full sense to perceive it at all.

3.8.4 Apperception, self-control, and education

Apperception is closely connected to morality, as we will see in more detail in the section on Herbart's ethics. It is interesting to note, however, that even in his discussion of inner perception (or as we might call it, "self-awareness"), his examples all have a normative dimension. Thus, as in the case of an unattended *external* humming noise, which we suddenly realize has been going on in the margin of consciousness, so too we most easily recognize *inner* perception when it lapses, e.g., when one lets slip a secret, or inappropriately laughs or yawns (SW VI: 145). One *ought* to have noticed and resisted the incipient impulses towards these actions (SW VI: 145). In short, *self-control* is a case of inner perception. Hence, in the case of "weak or uneducated people, children, and esp. of animals", one does not "expect them to mark their inner conditions" (SW VI: 145).

On Herbart's theory, such cases of failing to arrest or "fix" a representational evolution may be explained by the absence of representational masses sufficiently articulated and strengthened, not just into "concepts", but into clear judgments formed out of these concepts, which are called "maxims" (SW VI: 146). It is the *educated* person who "should know better", who is expected to have maxims that arise to ward off foolish, inappropriate, or unethical impulses. These "maxim"-masses are called "inner perceptions" (self-consciousness or self-control) just when they successfully "outstrip, fix, and govern [*beherrschen*]" rough and precipitate excitations (SW VI: 146).

The German word for "education" is "*Bildung*", literally: "formation". Now, apperception depends on the action of dominant representational masses, and self-control and self-awareness are functions of apperception. Thus the educator has a keen interest in forming or shaping the incipient concept- and maxim-masses in the young. Unless raw masses are strengthened and clarified, they will remain "mere heaps [of representations] lacking inner formation and order [*Ausbildung und Anordnung*], as in the case of rough persons" (SW VI: 147). The more reticulated representational chains become, the more defined, stable, and effective are the resulting concept-masses. These in turn allow for the most acute perception, as a

> skilled musician … will pick out of a large choir of voices the one which is at fault … [or the] physician perceives in an instant symptoms which have escaped the long and anxious scrutiny

of the inexpert (Stout 1888b: 486). Similarly, the development of moral concepts will produce the morally perceptive person, as we will discuss in Section 5.

3.8.5 Apperception and ego-consciousness

Apperception or "self-consciousness" seems naturally to involve the notion of the ego or "I": for what could self-consciousness be other than a self or ego that has somehow turned its thought back upon itself, so as to be simultaneously subject and object of awareness? (SW VI: 169) Now Herbart believes that his account of apperception will reveal the substantial "I" to be but an illusion generated by the interplay among differential representational masses (SW VI: 141). Nevertheless, although there is no such *thing* as the "ego", yet this does not mean that there is no *phenomenon* answering to the word, "I". We constantly make use of the first-person singular pronoun, which somehow expresses a first-person awareness, an *ego-consciousness*, and this psychological state is clearly of central importance. Herbart tries to explain the ego (correctly conceived) in naturalistic terms, viz. as the evolution of representational masses from childhood to adulthood.

He analyzes the genesis of self-consciousness as the evolution of apperception. From a complex of bodily sensations, desires, and aversions, the child becomes aware of representations in *other* beings, which it conceives as pictures carried around in their bodies. As that conception joins with observed intentional activities in animals and people—i.e., actions they carry out for their own satisfaction—the "itself" arises, though in the third person. Only when such recursive phenomena are apperceived within my own body does the representation of *my* self arise (SW VI: 177). This bodily self, then, appears as the "gathering place" not just of my representations, but also as the "*origin of all movements* connected physiologically with bodily feelings", the pole of all attraction and repulsion.

3.8.5.1 The growth of ego-consciousness

This nexus persists through time, and constantly receiving additions, forming and reforming itself. Hence, *Bildung*, in the most literal sense, is the constant accumulation of *Bilder* (pictures) around this node (SW VI: 177). As the human being begins to reflect upon its "self", it becomes aware of its overriding determination as a represent*ing*, know*ing*, cogniz*ing* entity; and the further its *Bildung* advances, the more this conviction grows (SW VI: 178; 180). Hence, "I", the being that exists in this way, become aware of myself as not a thing, but an act-*ing*.

Especially important for the *Ausbildung* or development of self-consciousness is the reminiscence of the past. For in this way one comes to see a difference between that center of *past* (psychic) conditions and current ones; hence all psychic conditions are recognized as *accidental*: the individual itself cannot be reduced to any one of them, and hence appears, in apperception, to stand forth as *distinct* from them (SW VI: 179). Thus, we come increasingly to see our selves as distinct from our external environment, and from the here and now, and a (representational) complex emerges "of which all the component parts may be negated, so that none of them seems essential to it"—and this is the ego (SW VI: 180).

Although the ego comes into focus through reflection on its own past, yet it is always in the present, projecting itself into the future. This leads to the conception of the ego as a "drive" (SW VI: 182). But a drive *to* sense, experience, or act always implies a goal. Hence "the ego appears most clearly in its outer activity" (SW VI: 182): it *represents* its action to itself, and assumes a cognitive distance from its actions. "It knows what it is about to do, and it also knows what it did" (SW VI: 182). In this way, the ego progressively abstracts itself from its activity, and eventually appears to itself as a "mere knowing that now has no object" (SW VI: 182). Further, by being with others in public, we come to notice an inner private sphere; for while other people and things may appear opaque, "the ego is always present to itself" (SW VI: 183). Thus the educated (*gebildet*) person comes to ascribe

> a soul to himself, indeed a *character*, [which he respects] as higher than the body. The inner perception is [now] seen as the source of knowledge of the true self. (SW VI: 186)

3.8.5.2 Dispelling the illusion of the ego

Herbart continues:

> But now comes philosophical reflection, which denies true self-knowledge to inner perception. It does not care about the *temporal being*, the *individual*, but wants to know about its [*sc. the individual's*] enduring and fixed foundation. It now discovers itself gradually, viz. that perception utterly lacks the actual soul-essence, the substance of the soul; and that such a substance has to be *attached to it in thought*.... Nevertheless, the ego remains the genuine, ever selfsame identity of the representing with the represented [thing]. This ego appears as a given, as the most secure and indisputable fact of consciousness. ... A peculiar kind of cognition is invented for this: a *pure, intellectual power*. But if one asks *what this [thing] is*, which intellectual intuition intuits, there appears the absurd concept of an ego severed from the individual. (SW VI: 186; cf. Beiser 2014: 133)

Thus, the ego reaches the complete illusion of itself as a perduring substance.

3.8.5.3 The method of relations, again

Herbart corrects this metaphysical illusion of the ego-soul through his method of relations.

> As with all concepts that lack an essential component [yet] to which they *stand in relation*,
> without *containing* it or immediately indicating it: so here, too, in the case of the ego, the illusion lies in this, that one imagines this concept to be thinkable [even] after separating out everything particular and individual. (SW VI: 187)

In other words, abstracted from its "contents", the ego does not exist: it

> is a point, that is and can be represented only insofar as innumerable [representational] sequences point back to it, viz. as their common presupposition

or anchor (SW VI: 168). It is merely a notional point of reference, a common node of all representational series (SW VI: 168).

This analysis brings into focus the relation of soul to ego. Recall the metaphysical axiom, that the soul is a simple whole—originally *not* representing—whose self-preservations against various disturbances result in "represent*ing*s". The soul in itself, in its simple and thus unknowable quality as *not representing* soul, cannot become either subject or object of consciousness. But the soul *with respect to* all of its representations is the true subject: the one, undivided, yet most variously active subject of all consciousness. Now as we consider psychic activity to be *generally* a represent*ing* over against a represent*ed*, there arises the peculiar case where the represented entity is identical with itself (SW VI: 190–1). The soul "unfold[s] the represented world; and, in the middle of the world, the *represented*, own self" (SW VI: 191).

This *represented* self, then, is the ego, the apperceptive system of the soul's perturbations. As we will see in the final section on Herbart's pedagogy, education just amounts to the best development of this ego, i.e., its virtuous character. This education (*Ausbildung*) culminates in science, by which Herbart means the soul's self-knowledge in terms of *his psychology*. For this "*science speaks of the soul* as the ground of the represented world and of its own Self". Thus, while Herbart's language of a culminating, reflective science of the self may remind one of an Idealistic absolute science, he takes great pains to distance his notion of science from *speculative* or *intuitionistic metaphysics*. Yes, in "science", the knower is the soul, and "knowing and known are one and the same, viz. the soul in the system of its self-preservations"; however, he adds, "I know of myself, not with innate, but with a permanently *acquired* knowledge" (SW VI: 191).

4. Aesthetics & ethics

4.1 Aesthetics

Whereas metaphysics and psychology deal with the real, aesthetics concerns just those concepts "that solely indicate a value or disvalue" (Weiss 1928: 108; cf. SW XI: 24; 26). Aesthetics is thus a *science of values*. In particular, aesthetics inquires into the beautiful, which is not a cognitive concept, i.e., no predicate of beings insofar as they exist. Rather the beautiful "denotes the way ... an object is apprehended by an observer..." (SW IV: 500; cf. Weiss 1928: 109).

Beauty lies in the eye of the beholder, and is an *experience* of the subject. At the same time, the beautiful is not something merely subjective like a feeling of pleasantness, for it arises and subsists in the observer only *with respect to* an object (cf. SW IV: 602). What the aesthetic conception articulates about the object does not "depend on any individuality, but rather solely on the quality of the represented [object]" (SW VI: 273). Suppose I declare: "Rembrandt's *Aristotle with a Bust of Homer* is beautiful". On the one hand, the beauty of the painting is something that lies not *in* the painting, but in my, the beholder's experience of it. But the predicate, "is beautiful", does not attach to my experience, but to its object, the painting. I do not claim that my experience is beautiful, but that the painting is. Nor am I merely saying that it *seems* beautiful *to me*; rather, the claim pertains to the painting itself, viz. that it *is* beautiful.

4.1.1 Aesthetic judgment

For Herbart, an aesthetic judgment is marked by the *immediacy and involuntariness* with which it is made. Such judgments are, in other words, independent of the personal preference of the subject, and are, in a sense, forced upon her. They are warranted by a peculiar, originary kind of evidence springing from the observation of the object (SW X: 318). They are clear, without having been learned or proved, and stand forth as absolute and non-negotiable claims (SW IV: 105; 47). This absolute and self-evident character lets aesthetic judgments function as the principles of a *science* of aesthetics.

> They let aesthetic feeling be brought into the form of a judgment, let the beautiful be treated scientifically, and let aesthetics become a science of principles. (Weiss 1928: 109)

Such immediate, involuntary judgments arise upon the "complete [or perfect] representation of the object [of the judgment]": "the beautiful is the effect of complete [or perfect] representation" (Weiss 1928: 110). By "complete representation", Herbart means the kind that arises of itself in the utterly non-partisan observer, moved by the object alone. A beautiful object will necessarily find favor, an ugly object distaste, for it appears clearly to one free "of the disturbing pressures and demands of the will" (SW III: 526). In short, aesthetic judgment depends on a contemplative state of ataraxic indifference or will-less-ness (*Willenlosigkeit*) (cf. SW II: 335; Beiser 2014: 126).

4.1.2 Aesthetic validity and "conscience"

The aesthetic judgment is synthetic, connecting the theoretical cognition of the object with its evaluation. The former is the subject of the judgment, the latter its predicate (SW IX: 80; 329, n.*). Herbart does not, of course, regard such judgments as functions of a *power* of judgment. Rather, judgments of taste are the involuntary effect of any perfect representation, "provided the latter does not immediately disappear in the flux" of psychic change (SW II: 45). The aesthetic judgment seems to possess an absolute validity. This might seem strange, given the subjectivity of taste, but as we saw above, the neutral judge's judgment makes a claim about the object (cf. Kant, *Critique of Judgment*, §§2, 4). But there is a further problem: if each aesthetic judgment pertains to a single and particular representation, how can it have universal validity? (Cf. esp. Beiser 2014: 129) On Herbart's view, it gains the "appearance of an enduring, indeed eternal authority", thanks to a kind of "subjective necessity" (SW II: 350). That is, although each new perfected representation inevitably draws after it a new judgment, yet it is also true that the same representation will always provoke the same judgment (SW II: 350). In other words, under identical subjective and objective conditions, the same aesthetic judgment *must* recur, and this produces the appearance of eternal validity.

Now, as one repeatedly passes aesthetic judgments and becomes aware of their peculiar validity, they come to exert "a slow pressure upon the [individual], who calls it his 'conscience'" (SW I: 265). Only he who repeatedly generates aesthetic judgments in fact acquires such a "conscience". We may take Herbart to be speaking here of a sense of artistic propriety characteristic of the connoisseur or the trained musician. In musical practice, for example, one makes myriad judgments regarding the tempo, articulation, or dynamics. Over time, these fuse into stable appercipient masses—the musical "conscience"—and music conflicting with them seems "wrong" and *must* be criticized. Such a Herbartian conscience is no kind of faculty of taste or judgment, but a facility acquired in accordance with the laws of concept formation.

4.1.3 The beautiful

Much as Aristotle rejects Plato's unitary conception of the Good, Herbart emphasizes the particularity of the beautiful. Aesthetics cannot and ought not try to abstract from individual judgments of taste in order to discover "the" common essence of the beautiful (or the ugly). This is because Herbart holds the original aesthetic judgments, *in their absoluteness*, to be the principles of aesthetics.

> [Their] evidence, … absoluteness, and … perfect representation of the object as the indispensable preconditions of the possibility [of such judgments] exclude the possibility of any highest aesthetic propositions [*Grundsätze*] … to which a number of aesthetic judgments … should be subordinated. (SW IV: 47)

Although it *is* universally true of the aesthetic that the object involuntarily pleases or repels, yet the objects of aesthetic judgment themselves cannot be reduced to a formula. If one abstracts from their particularity, one is left with "no object containing something that an aesthetic judgment could determine".

What is it in the object that elicits approval or disapproval—what lets an object be judged "beautiful"? Beauty is a function of relations (*Verhältnisse*) or form, as in the "relations in the outline of a statue, in the characters of a drama, of tones in musical chords" (SW IX: 329; cf. SW IX: 313, SW II: 344). For this reason, Herbart holds that only an object exhibiting internal structure or a relation to other objects can have aesthetic interest. An object without relations is "aesthetically insignificant" (Weiss 1928: 115). In music, no single tone can be considered, in itself, beautiful or ugly, although single tones constitute the intervals and chords that, as relations, elicit aesthetic judgment (cf. SW II: 344). Hence the axiom of all aesthetics is this:

> All simple aesthetic elements must themselves be relations, viz., relations whose *relata*, considered in themselves, have no aesthetic value. (SW III: 331)

We now see better what it is that aesthetic judgments express approval or disapproval of, viz. relations that can be grasped in a single representing act (Weiss 1928: 116).

4.1.4 The task of aesthetics as a discipline

Thus aesthetics starts with the simplest formal determinations underlying approbation or disapproval (Weiss 1928: 117). Its disciplinary task is

> neither to define, nor demonstrate, nor deduce, nor even to distinguish amongst types of art or argue about a given work of art. Rather, it should reveal the *simple relations* that generate approbation or deprecation in case of a perfect representing. It should make us *conscious* … of the specific approbation and of the specific disapproval that is originally proper to any particular relationship whatsoever. (SW II: 345)

The concepts we use to think these elemental relations Herbart calls "paradigms or Ideas" (Weiss 1928: 118). The exhaustive presentation of these elementary relations grounds aesthetics as a science. Once they are in hand, it can survey all possible aesthetic relations, as in figured bass. Indeed, Herbart considers figured bass to be the only existing model of an aesthetic science, for it "demands and receives an absolute judgment for its simple intervals, chords, and progressions, without proving or explaining anything" (SW II: 345).

Should these universal aesthetic Ideas be organized in a general aesthetics, they could then be applied in different material spheres, e.g., of color, tone, figure, but also, crucially, of *relations of the will*. Such specific ramifications of the paradigms are called "techniques" (*Kunstlehren*) or "practical sciences" (SW IV: 47). The discipline dealing with the paradigmatic elements of approbation or disapprobation of relations of will is *practical philosophy*, i.e., ethics, to which we now turn.

4.2 Ethics

All evaluative judgments are aesthetic; they evaluate the represented relations in their object. *Practical philosophy* or ethics evaluates relations of the will. Therefore, ethical judgments are a kind of aesthetic judgment (cf. SW VII: 10; SW X: 309; Beiser 2014: 108; 125, ff.). Because such judgments often conflict, the task of practical philosophy is not to pass judgment itself, but to cause

us to judge correctly. Hence, all practical philosophy can and should do is to "correctly present the object, [in such a way] as to be perfectly grasped", that object being relations of the will (SW II: 334); thus, as figured bass presents a theory of beautiful tonal relations, practical philosophy seeks the "whole set of *relations* by which the will becomes the object of praise or blame" (SW IX: 329: n.*). Like Kant, Herbart holds that the determining grounds of moral action must be found in *a priori* norms, and rejects eudemonism or any "material motives" as grounds for practical laws (cf. Beiser 2014: 94, 125, ff.). For Herbart as for Kant, morally evaluative judgments are to be made solely with an eye to the "form of the will", which ethics determines.

4.2.1 The object of practical judgment

Practical judgment judges the "will", a species of desire marked by determinate cognition and fixing of its object, and by conviction that the object is attainable. While I might desire to leap tall buildings in a bound, I cannot will it (SW II: 99). Now the value of the will's object cannot determine its evaluation, since Herbart like Kant excludes all material grounds as determinative of moral action. Hence the will itself is the object of evaluation—but how? Because every instance of willing is nonetheless bound to its object, *individual* acts of willing cannot be the object of moral evaluation. Further, an act of willing isolated from its object also cannot be the object of moral evaluation, since

> otherwise, the difference between a good and evil will would disappear: willing and willing-well would be one and the same. (Weiss 1928: 122)

Herbart's solution to the dilemma is to say that practical philosophy considers the will not in itself, but only insofar as it enters into relation with another will.[68] It is these various possible relations *among wills* that constitute the possible forms of the will, and that are judged fine or shameful by practical judgment (SW X: 348). In this way, the aesthetic elements and paradigmatic Ideas of general aesthetics govern practical philosophy, as well.

4.2.2 The paradigms or Ideas of practical philosophy

Herbart identifies five such elementary relations of will that demand aesthetic approval, viz., the Ideas of inner freedom; perfection; benevolence; right; fairness/justice (SW II: 356). The first Idea is the Idea *ot* Ideas, i.e., the general decision or determination of the will always "to act in accordance with the patterns *as* patterns, in accordance with taste in general" (SW II: 357). Unlike the other Ideas, inner freedom is a paradigm that is unconditionally pleasing, and which therefore does not need to be commanded. On the other hand, the will can easily come into conflict with the other four paradigms, and such conflict is *displeasing*. The insight of practical judgment into the paradigmatic pleasing-ness of the Ideas therefore makes itself felt as a demand to neutralize that conflict:

> [t]he ethical commandment springs solely out of judgments of displeasure, and appears when, and for as long as, a willing conflicts with the practical Idea.

4.2.3 The virtuous human being

Herbart's Doctrine of Ideas naturally leads to the notion of virtue or human excellence, for "the ideal, complete realization of inner freedom [just] is the *virtue* of the individual" (Weiss 1928: 145; cf. SW II: 409, ff., esp. 411–12). Here we advance from aesthetic pleasure in the paradigms of the will to their actual realization: "Virtue does not forever stay home but strides forth into an alien world" (SW II: 414). Thus the individual who sees and approves of the practical Ideas also feels compelled to enact them in a coherent way of life. This unity is the *Ideal of virtue*, according to which "no part of [practical] insight may lack the corresponding will[,] and there should be no lack of connection among its parts" (SW II: 412). The virtuous life must exclude vice and lack of virtue, and react to the situations that at every turn demand virtue's activity (SW II: 412–13).

Instead of an indefinitely ramifying list of duties to achieve this ideal, Herbart proposes that the sense of art or *Kunstsinn* can help us gain an overview of virtue (SW II: 415). That is, life's circumstances appear as a matter to be formed, and the virtuous person is able to judge the *finest* way in which a situation should be handled (SW II: 416). But because the artistic sense is in constant conflict with external forces and internal feelings, it must be continually rekindled (SW II: 417). Hence, the Ideal of virtue can be at best asymptotically approached and it is this approach that Herbart calls "morality". Since an approach implies a temporal sequence, human life may be understood as a chain of ethical actions and passions (SW II: 417, f.).

Herbart analyzes this chain into three main parts. First, there is virtue as "the picture of a beautiful reality", belonging to imagination, dream, and fantasy, in which the need for intervening in reality is lacking: the Ideas are simply assumed as universally effective (SW II: 417–18). Second, the need to act arises, and the striving will encounters obstacles even within the sphere of thought. Finally, we really act, and try to realize the beautiful world of thought (SW II: 418). While social reality hampers the achievement of virtue, the individual, conscious of his inner freedom, must try to find all possible avenues towards moral fulfillment (Weiss 1928: 156). It is in preparing the individual for a life of moral action that education is required.

5. Pedagogy

Herbart writes:

> The sole and entire task of education [*Erziehung*] is
> encompassed by the concept, "morality" [*Moralität*].
> (SW I: 259)

If virtue is the realization of human freedom, then the goal of Herbartian education is to make the child *capable* of living virtuously, viz., by forming the pupil's character. For reasons that will become clearer below, Herbart identifies moral character with a "many-sided" character. Many-sidedness is a regulative ideal guiding one's striving towards full individuality: the teacher helps the pupil develop all his[69] unique talents and interests as far and as coherently as possible. That is, the teacher helps the pupil systematically connect his interests around the central principle that is *his self*.[70]

5.1 Many-sidedness and "interest"

Instruction aiming at many-sidedness has two main phases, "engrossment" (*Vertiefung*) and "reflection" (*Besinnung*), each of which in turn has a "resting" and a "progressive" moment: resting engrossment is "clarity"; progressive engrossment is "association"; resting reflection is "system"; progressive reflection is "method" (SW II: 38, ff.; cf. esp. Glöckner 1892; Beiser 2014: 106–7). It is important to note that these phases of learning recapitulate the process of apperception. That is, clarity occurs when a representation or representational series comes into full consciousness, where it hovers in the full light of attention. One is engrossed *in* it, which is simply to say, one is not aware of anything else. But of course the child's mind rarely stays occupied with a single representation for long; its imagination (*Phantasie*) rapidly leads it, usually by association, to other, similar representations (SW II: 40). This spontaneity of absorption in environing objects is the source of a *natural* many-sidedness, and thus the seed of the fully formed personality. While the phases of engrossment are natural to all human beings, it is obvious that in themselves they have no direction or organizing principle.[71]

The teacher does not simply give the principle to the child. Rather, his instruction aims to bring the child by stages to recognize the principle on and as its own, by *reflecting* on its immediate experience. The child's maturity determines the scope of reflection, but she will always fall short of the full self-reflection characteristic of the adult. Since many-sidedness depends on reflection for its full solidity, the child can also only ever approximate adult versatility and virtue (PS I: 149). For Herbart, the teacher promotes reflection by promoting what he calls the child's "interest" (*Interesse*). Now, the child's natural engrossment in its activities must be respected as symbols of its individuality. At the same time, engrossment is a kind of seclusion, whereas interest is an opening up of the person in and to the world.[72] Step by step, the teacher must lead the pupil to become an interested, active learner, who desires not to possess knowledge, but to hover "*between* merely spectating [on the one hand] and grasping [on the other]" (cf. SW II: 42).

The first phase of interest is *noticing* objects:

> the noticed object not only stands forth for a certain
> duration from the flow of psychic activity, but also aids
> the reproduction of suppressed representations. (Weiss
> 1928: 194; cf. SW II: 42; 53)

As other representations rise, new, imagined representations are anticipated (as we saw frequently happens in apperception). Should an imagined representation be perfectly represented, it will naturally provoke an aesthetic judgment of approval. This leads to a "demand" for the imagined object, i.e., an overpowering sense for its realization; and this demand finally expresses itself in action in the world (SW II: 42–3). It is thus clear that by "cultivating interest", Herbart does not mean simply getting the pupil "to be interested in something". Rather, as these four phases of interest map onto the phases of engrossment and reflection, interest is an expression of many-sidedness, namely its activity in the world. For "demand" here is not simply a strong desire, but an apperceptive "mental anticipation" of the object's realization. As such it is necessary for the pupil to advance to a more perfect state of systematic and aesthetic completeness.[73] Thus by promoting interest and many-sidedness, education primes the pupil for action (PS: 181).[74]

Thus far, "interest" has remained quite abstract: it is, *in nuce*, a questing openness towards certain objects, leading to action. Now the kind of action Herbart has in mind is not an instinctive reaction, but is integrated into a systematic plan. For the totality of "the interesting" is prescribed by *culture*,[75] and its divisions into the cultural and natural sciences (SW II: 44, f.). Education must accordingly pursue in parallel the formation of the pupil's imagination and understanding (PS I: 59; cf. 181). Because all knowledge is rooted in experience, organized through speculation, and grasped as a whole aesthetically, *Bildung* must develop these three dimensions equally (SW II: 53). Equally important, however, is the pupil's constant, active participation in his surroundings, in order to develop an empathic understanding of her fellow human beings (PS I: 156; SW II: 51, ff.). In all these cases, a teacher is needed to modify experience and human interaction in order to maximize the yield of knowledge. For experience merely accumulates random representations; *instruction* helps the

person to analyze, organize and combine these "formless fragments" (Weiss 1928: 200).[76] By the same token, mere everyday contact with others does not produce empathy.

Hence, the teacher must not be narrowly committed to his scientific expertise, but demonstrate interest in the world and human beings (SW II: 50). While native taste and thoughtfulness cannot be taught, the teacher must be alert to these and fan such sparks in order to *complete* what nature has prepared (SW II: 46). Further, he must respect the "personal unity" of the child: Herbart's rule:

> to give equal right to engrossment and reflection in every in area of inquiry, no matter how small; i.e., to tend in order to: *clarity* of the particular; *association* of the many*; coordination* of what has been associated, as well as practice in progressing *through* this organization.

In all this, the elements of instruction—"facts, information, ideas, knowledge"—are not negligible, but rather

> vitally important, for … they become built up into "apperception masses", which, in the process of taking in *more* facts, information, ideas, knowledge, give rise to "apperceptive interest", this latter being itself of first importance in the character-forming process. (Hayward 1907: 70)

5.2 Character

We have so far been discussing the ideal of the many-sidedness; the many-sided person and the person of character is one and the same, though regarded from different perspectives (PS I: 199; SW II: 89; cf. SW IV, esp. "*Pädagogisches Gutachten usw.*": 519, ff). Herbart is particularly concerned with *moral character*, and its virtues of stability, hardness (*Härte*), and invulnerability. He wants to show that a many-sided interest serves moral character, and how pedagogy devoted to the former contributes to moral education (Weiss 1928: 211). For Herbart, (moral) character, like all spiritual (*geistig*) aspects of human life, is not a natural given, but a gradual acquisition. Psychologically speaking, it is the will that in the young child is stamped through repeated interaction with the outer world, settling into "uniform and stable [ways] in which [it pursues] certain goals while excluding others"; thus, the will becomes *character* (Weiss 1928: 211; cf. SW II: 90).

Character has distinct objective and subjective sides (SW II: 90–1). Objectively, character depends on the child's susceptibility to formation and its individuality. Although the child is not born with "a" character (which term is reserved for settled traits), yet he is born with certain dispositions, including a "disposition to stability of character" (SW II: 91). This manifests itself in what Herbart calls the will's "memory", i.e., the fact that under like stimuli, the will will will like actions (SW II: 91–2). In this respect, there seems to be something predictable, hence "objective" in human character. Although this stability is but relative, and one spontaneously finds himself choosing or forsaking, avoiding this or pursuing that, yet

> where there is a memory of the will, the choice [in a given situation] will be made of itself. The preponderance of wishes will involuntarily organize itself [and manifests itself in] choice. (SW II: 92; SW V: 203)

In different individuals, their inclinations and naturally differing intensity will cause the stability of simple willing to vary: "every individuality is and remains a chameleon" (SW II: 93). Nonetheless, they originate in a "unified source and point of reference" that constitutes the "essential core of individuality" (Weiss 1928: 213). Hence while choice only occurs in accordance with an individual, particular standard, it remains the precondition for the consolidation of character itself. It is out of simple willing and the psychological character to which it gradually gives rise that the "matter" is gathered, as it were, with which one has to work, and from which he ultimately "must secure the demands that he makes of himself": "Being is what gets a person started; demands follow after" (PS I: 198).

On the other hand, these demands typify the higher, subjective aspect of character, "higher" because marked by reflective intelligence, an apperceptive distinction between willing and the objects of will, between I and not-I. Thus the ego now identifies the natural or simple will as its "own" will: the natural stability of the simple will becomes a willed or conscious stability. In addition, the reflective effort to grasp oneself leads to a *demand* regarding the objects towards which *conscious* willing is directed. That is, the will must adapt itself to changes in the object: it perceives a *demand* that one must "precisely orient oneself in view of circumstances", and to do so differently as circumstances

shift. Hence the stability of character must be tempered with flexibility (*Motivität*); in the "precision of this flexibility lies the art of *characterful* living" (PS I: 201; 351, f.; cf. esp. 198). Individual willing settles into character through the struggle that arises in the individual's attempt to defend the ego and the will's stability. Because the stability of character is only relative, and shifts over the years, its true *principle* can only consist in being oriented towards a realm of purposes (*Zwecke*) or ultimate values (*Werte*) (Weiss 1928: 216).

5.3 Ethical aspect of character

> To bring it about that the pupil finds
> himself, *as* choosing the good, *as* rejecting the evil: this
> —or nothing—is character-formation! (SW I: 261)

Character has an ineluctably ethical aspect: we want (our) character not to be of just any sort, but good. Like Aristotle, Herbart holds that we always stand in some relation to virtue and vice. While not by nature good or bad, our actions and character necessarily fall somewhere on the spectrum of virtue and vice. Virtue is thus not something attached from the outside to an ethically neutral self; rather, it is a natural quality of man that will develop well or poorly. So Herbart, like Aristotle and Plato, does not think that education should aim to inculcate moral truths or commandments, so much as lead the "moral[ly constituted human being] to *recognize* his own value and dignity, and [then] to elevate it" (PS I: 199). Here again we see the transition from the objective facticity of one's constitution as a being with *some* moral charge, to the subjective endorsement of it, of accepting responsibility for the excellence of one's moral constitution as the source of one's value and dignity. It is a movement of "awakening self-reflection", in which the form-law governing our involuntary (aesthetic) approval or disapproval of human actions or situations is elevated to norm-law, i.e., is accepted as binding on me, and to which I have a responsibility that I accept (cf. Weiss 1928: 216).

As with character generally, moral character displays a kind of natural memory, a "perduring in simple willing [,] the natural beginning of unity and consistency of moral willing", already evident in the small child's incipient moral feeling (PS I: 199; cf. SW II: 120–1). All this points to the fact that in the objective or natural constitution of character, we already find "conceptions of the good and right". This natural moral memory shows itself in the fact that

> moral judgment as the pure natural appearance of the
> essence of humanity is not "obscure" and "mutable",
> but [always] "decisive", that it always appears as the
> same willing whenever the same stimulus is renewed
> and nothing hinders or neutralizes a pure conception of
> the object, [i.e.,] whenever the mind is in a state of pure
> contemplation. (Weiss 1928: 217; cf. esp. PS I: 131)

This consistent, disinterested willing just is the phenomenon of conscience, which thus reveals itself as a form of willing, namely the natural desire for the good as neutrally judged. As will, conscience, too, has an object, namely the moral ideal, i.e., its own relation to other wills. As we saw in our discussion of Herbart's ethics,

> [t]he judged [thing, sc. judged by "conscience"] is also
> [the] will, [only in this case] the will in its relation to
> other wills (PS I: 202)

Conscience as "natural" or "objective" will thus articulates the (moral) demands that the human being makes of himself, the pursuit of which determines his moral character, now properly so called. Self-reflection at first notices objective as well as subjective willing, but when the object has been completely represented, viz., the individual will as a member or *relatum* of a will-relation, then this complete representation of the object also provokes approval or disapproval as an immediate "aesthetic" effect . In other words, the individual goes beyond merely registering how he feels or finds himself, to positively approving or censuring that same inner state.

We see here a special case of apperception as an interaction among conflicting representational masses, in which one representational group is not only "noticed" but evaluated by another, held to what (from its perspective) is perceived as an alien claim. Now, on the one hand, since the soul is postulated as a unity, the difference between how one *is*, and how one judges that one *ought to be*, indicates an inner conflict. On the other hand, the moral personality, as a manifestation of the integral soul, must or ought also to be integrated. In short, the conflict must be resolved so that the judged will is brought into agreement with the aesthetic (apperceiving, evaluating) will, and so "merge into a single person": "[T]his person represents the ideal of an ethically self-determining will" (PS I: 202; cf. 131). But since this ideal is rarely if ever achieved, we constantly feel compelled to "interrupt our activity and alter our mental disposition from within". Thus, for the sake of the unity of the moral personality, the human takes as a categorical imperative what previously was but an aesthetic judgment.

Now it is a presupposition of pedagogy that all human beings are capable of living up to the ideal of an integrated, unified, autonomous personality (PS I: 202; 93). But how does the will become decisive? In other words, how do we gain a moral character that *acts* for the good? We may approach this problem by asking how those natures that successfully form themselves do so, and through what stages they must pass in "gaining a character" (Weiss 1928: 221–2). This can help the pedagogue develop a plan for optimizing what is otherwise a haphazard process. The resolute character depends, first, upon a moral *eros*, and striving towards a representation. The will generates the act out of a desire (PS I: 360). But active deeds in turn depend on natural disposition—including bodily health; the will's memory; and the spirit's mobility or "flexibility" (PS I: 365). The less motile the spirit, the firmer the will (PS I: 365; 198).[77] Now natural disposition develops very slowly and only fully matures in adults when they enter the world. Hence character-education has little latitude, since children rarely are placed in situations of true responsibility. At just the moment when the young adult takes up responsibilities, his receptivity for education is exhausted (SW II: 88, f.; PS I: 364).

Herbart draws two conclusions: First, that character-formation must focus on the young person's "sphere of desire [*Begehrungskreis*]", as the one aspect of character apparently susceptible to "a lasting influence" (Weiss 1928: 224). And since desires aim at goals, and goals are representations, the educator can set to work upon the "sphere of representations", and thus shape the ultimate targets for the mature individual's future desire (SW II: 118, f.). Second, Herbart concludes that character-development gains stability and security by being accelerated and integrated into the actual period and process of education. The teacher must keep in mind that character "is solely formed through action out of *one's own will*" (PS I: 251). Therefore, Herbart's pedagogy emphasizes the importance of offering the child's will ample opportunity to express itself: "Boys and youths must be *dared* to become men" ("*Maßregeln der Kinderregierung*"; Herbart 1887b: 17). It is the work itself that fixes the bounds of thought and action, not the teacher; within those bounds the student is to be given maximal freedom (PS I: 308). By frequently exercising the freedom of the will in purposeful activity, the pedagogue is able to "harden" the pupil's will (PS I: 368–9).

We began by asking how moral character is to be formed. This question finally calls for consideration of the "influences that specially affect the moral features of character" (Weiss 1928: 225). Chief among these is aesthetic judgment, which must therefore be constantly exercised, so as to become the power mastering desires. The pedagogue must therefore carefully attend to the contents of the child's or sphere of thoughts. First, the pupil should gain a clear and particularized familiarity with a "whole series of moral elements and their most common causes in life", e.g., the main relationships of daily life, family, country, and their various conflicts and contradictions. Further, since character ultimately depends on the integration of the person, all domains of human interest must be brought into relation with the ideal: thus, for Herbart as for Montaigne,[78] history in the broadest sense provides an indispensable space in which to exercise the young person's aesthetic and practical judgment. Since passive presentation of historical, literary, or philosophical *pictures* must be avoided, the instructor must strive to bring peoples and persons of the past to the fore *as living people*: practice in recognizing many human types sharpens the pupil's "ethical eye", securing the child from dangerous surprises in life (SW II: 50). Again, because historical figures appear at a distance, the pupil can contemplate them in a disinterested fashion, letting "pure taste" rule action in his imagination. Such training is the best one can do in preparing the person for adulthood.